

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 18:57:59 ; Search time 75 Seconds
(without alignments)
2140.173 Million cell updates/sec

Title: US-10-051-909-32

Perfect score: 4075

Sequence: 1 IRSGSWLAVQPTFPDRLRR.....PLEVITEFVAGAKQAANKA 800

Scoring table:

BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 733937 seqs, 200641211 residues

Total number of hits satisfying chosen parameters: 733937

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09D_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4075	100.0	800	US-10-051-909-32	Sequence 32, Appl
2	3774	92.6	745	US-10-310-154-718	Sequence 718, Appl
3	3517	86.3	747	US-10-051-902-2	Sequence 2, Appl
4	3517	86.3	747	US-10-051-909-2	Sequence 2, Appl
5	2834	69.5	740	US-10-051-909-37	Sequence 37, Appl
6	2674	65.6	737	US-10-051-902-8	Sequence 8, Appl
7	2674	65.6	737	US-10-051-909-8	Sequence 8, Appl
8	2584	63.4	729	US-10-051-902-29	Sequence 29, Appl
9	2584	63.4	729	US-10-051-909-29	Sequence 29, Appl
10	1670.5	41.0	486	US-10-051-902-10	Sequence 10, Appl
11	1670.5	41.0	486	US-10-051-909-10	Sequence 10, Appl
12	1424	34.9	345	US-10-051-902-14	Sequence 14, Appl
13	1424	34.9	345	US-10-051-909-14	Sequence 14, Appl
14	943	23.1	228	US-10-051-902-16	Sequence 16, Appl
15	943	23.1	228	US-10-051-909-16	Sequence 16, Appl

16	587.5	14.4	457	12	US-10-369-493-23324	Sequence 23324, A
17	562	13.8	131	14	US-10-051-902-6	Sequence 6, Appl
18	562	13.8	131	14	US-10-051-909-6	Sequence 6, Appl
19	554.5	13.6	487	9	US-09-795-693-27	Sequence 27, Appl
20	554.5	13.6	487	12	US-10-391-399-45	Sequence 45, Appl
21	554.5	13.6	487	14	US-10-095-139-14	Sequence 14, Appl
22	554.5	13.6	487	15	US-10-156-239-27	Sequence 27, Appl
23	554.5	13.6	487	15	US-10-199-485-27	Sequence 27, Appl
24	548	13.4	488	12	US-10-162-102-46	Sequence 46, Appl
25	548	13.4	488	14	US-10-094-059-4	Sequence 4, Appl
26	548	13.4	488	15	US-10-170-528-5	Sequence 5, Appl
27	548	13.4	488	15	US-10-162-012-46	Sequence 46, Appl
28	548	13.4	488	15	US-10-062-9608-4	Sequence 4, Appl
29	548	13.4	488	15	US-10-144-624-4	Sequence 4, Appl
30	547	13.4	486	9	US-09-860-2324-7	Sequence 7, Appl
31	534	13.1	580	12	US-10-310-154-722	Sequence 722, Appl
32	533	13.1	461	12	US-10-369-493-23371	Sequence 23371, A
33	531.5	13.0	535	9	US-09-795-693-20	Sequence 20, Appl
34	531.5	13.0	535	15	US-10-156-239-20	Sequence 20, Appl
35	531.5	13.0	535	15	US-10-199-485-20	Sequence 20, Appl
36	505	12.4	473	12	US-10-369-493-23097	Sequence 23097, A
37	496	12.2	523	14	US-10-051-902-24	Sequence 24, Appl
38	496	12.2	523	14	US-10-051-909-24	Sequence 24, Appl
39	489.5	12.0	513	14	US-10-051-902-20	Sequence 20, Appl
40	489.5	12.0	513	14	US-10-051-909-20	Sequence 20, Appl
41	489	12.0	529	14	US-10-051-902-28	Sequence 28, Appl
42	489	12.0	529	14	US-10-051-909-28	Sequence 28, Appl
43	488	12.0	548	11	US-09-774-381-40	Sequence 40, Appl
44	488	12.0	549	14	US-10-051-902-30	Sequence 30, Appl
45	488	12.0	549	14	US-10-051-909-30	Sequence 30, Appl

ALIGNMENTS

RESULT 1
US-10-051-909-32
; Sequence 32, Application US/10051909
; Publication NO. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 32
; LENGTH: 800
; TYPE: PRT
; ORGANISM: Zea mays
US-10-051-909-32

Query Match 100.0%; Score 4075; DB 14; Length 800;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 800; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IRSGSWLAVQPTFPDRLRRRLPSVTLAPGLPAPSCSQSEPTSDTILEDMSGANV 60
DB 1 IRSGSWLAVQPTFPDRLRRRLPSVTLAPGLPAPSCSQSEPTSDTILEDMSGANV 60
QY 61 LVAIYASIGNLLOGMDNTAAVYIKKXQLONEPTVEGLIYMSLIGATVITTFSGP 120
DB 61 LVAIYASIGNLLOGMDNTAAVYIKKXQLONEPTVEGLIYMSLIGATVITTFSGP 120
QY 121 LSDSGRRPMLTSSILYFSGLITLWSPNYVLLAFVVGFGIGLAVTVPTVYISIA 180
DB 121 LSDSGRRPMLTSSILYFSGLITLWSPNYVLLAFVVGFGIGLAVTVPTVYISIA 180

121 LSDSIGRRPMLLSLTYFFSGILMLMSPNNVYLLARFVDGFGIGLATVLPVLYSEIA 180
181 PSEINGLNTLTPQSSGGMFLSYCVFQMSLSPSPDMRIMLVLAIPSLFFFGLTIFYL 240
181 PSEINGLNTLTPQSSGGMFLSYCVFQMSLSPSPDMRIMLVLAIPSLFFFGLTIFYL 240
241 PESPMLVSKGMAEKKVLOKLRGKDVSGELSLLEGEVGGDTISIEYIIGPTTEA 300
241 PESPMLVSKGMAEKKVLOKLRGKDVSGELSLLEGEVGGDTISIEYIIGPTTEA 300
301 DDLVTDGDKQITLYGPEEGOSWIAPSKGPIMLGSLVLSARHSGMNVQSPMDPIYT 360
301 DDLVTDGDKQITLYGPEEGOSWIAPSKGPIMLGSLVLSARHSGMNVQSPMDPIYT 360
361 LFSGVHNNPQAGSGSRSTLFPNFGSMFVTDHANNQWDEBNLHRDDEBYASDGAGD 420
361 LFSGVHNNPQAGSGSRSTLFPNFGSMFVTDHANNQWDEBNLHRDDEBYASDGAGD 420
421 YEDNLSPLSRQATGABGKDIVHGHRSALSMRQTLLEGGGDVSTDIQGGWQLAW 480
421 YEDNLSPLSRQATGABGKDIVHGHRSALSMRQTLLEGGGDVSTDIQGGWQLAW 480
481 KMSKEGNGRKEGKFRVYLHOGVPGSRGSIYSLPGGDPFEGSEFVHAALVSQSA 540
481 KMSKEGNGRKEGKFRVYLHOGVPGSRGSIYSLPGGDPFEGSEFVHAALVSQSA 540
541 LFSKGLAEPMSDAMVHPSEVAANKSRMKDLPEGVRALLVGVGIQILQOPAGINGVL 600
541 LFSKGLAEPMSDAMVHPSEVAANKSRMKDLPEGVRALLVGVGIQILQOPAGINGVL 600
601 YTPPOLBAGAVAILSKFSSASASLISLITLLMPCJGFALMLDLSGRRLLG 660
601 YTPPOLBAGAVAILSKFSSASASLISLITLLMPCJGFALMLDLSGRRLLG 660
661 TIFILIASVILVNSLIDGLTALHALLSTVSIVYFCFVWGFPIPNILCAIEPTRY 720
661 TIFILIASVILVNSLIDGLTALHALLSTVSIVYFCFVWGFPIPNILCAIEPTRY 720
721 RGLCTAICFTWIGDIYTYSLPWLNAIGLAFSIAVAVCLISFVFLKVPETKM 780
721 RGLCTAICFTWIGDIYTYSLPWLNAIGLAFSIAVAVCLISFVFLKVPETKM 780
781 PLEVITEFPAVGAQKAAKA 800
781 PLEVITEFPAVGAQKAAKA 800

RESULT 2
US-10-310-154-718
Sequence 718, Application US/10310154
Publication No. US20030233670A1
GENERAL INFORMATION:
APPLICANT: Edgerton, Michael D
APPLICANT: Chomet, Paul S.
APPLICANT: Adams, Thomas H
APPLICANT: Ruff, Thomas G.
APPLICANT: Agarwal, Ameeta K.
APPLICANT: Ahrens, Jeffrey E.
APPLICANT: Ball, James A.
APPLICANT: Banu, G.
APPLICANT: Bell, Erin
APPLICANT: Boddapalli, Raghava
APPLICANT: Delkman, Jili
APPLICANT: Deng, Molian
APPLICANT: Dong, Jinzhao
APPLICANT: Duff, Stephen M.
APPLICANT: Galligan, Meghan M.
APPLICANT: Hinchey, Brenda S.
APPLICANT: Huang, Shihshieh
APPLICANT: Johnson, G. Richard
APPLICANT: Jung, Vincent
APPLICANT: Kretzmer, Keith A.
APPLICANT: Laccetti, Lucille B.

APPLICANT: Lai, Chao-Qiang
APPLICANT: Lee, Gary
APPLICANT: Lin, Jie-Yi
APPLICANT: Liu, Jingdong
APPLICANT: Lu, Bin
APPLICANT: Luethy, Michael M.
APPLICANT: Lund, Adrian
APPLICANT: Madson, Linda L.
APPLICANT: Malloy, Kathleen A.
APPLICANT: McKel, Christine L.
APPLICANT: Miller, Philip W.
APPLICANT: Padmavathi, Manchi Kant
APPLICANT: Parnell, Laurence D.
APPLICANT: Start, William G.
APPLICANT: Tennessee, Dan
APPLICANT: Vidya, K. R.
APPLICANT: Wang, Haiyun
APPLICANT: Xin, Zhanqun
APPLICANT: Xu, Nanfei
APPLICANT: Yang, Chunzhi
APPLICANT: Zeng, Xiaoping
APPLICANT: Zhao, Yajuan
APPLICANT: Zhou, Li
TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
FILE REFERENCE: 38-15(52796)B
CURRENT APPLICATION NUMBER: US/10/310,154
CURRENT FILING DATE: 2002-12-04
PRIOR APPLICATION NUMBER: 60/337,358
PRIOR FILING DATE: 2001-12-04
NUMBER OF SEQ ID NOS: 736
SEQ ID NO 718
LENGTH: 745
TYPE: PRT
ORGANISM: Zea mays
US-10-310-154-718

Query Match 92.6%; Score 3774; DB 12; Length 745;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 742; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

56 MSGAVLVAIVASIGNLQGMNATIAAVLYIKKXEPOLNEPTVEGLVMSLIGATV 115
1 MSGAVLVAIVASIGNLQGMNATIAAVLYIKKXEPOLNEPTVEGLVMSLIGATV 60
116 TFSGLSDSIGRRPMLLSLTYFFSGILMLMSPNNVYLLARFVDGFGIGLATV 175
61 TFSGLSDSIGRRPMLLSLTYFFSGILMLMSPNNVYLLARFVDGFGIGLATV 120
176 ISEIAPSEIRGLNTLTPQSSGGMFLSYCVFQMSLSPSPDMRIMLVLAIPSLFFFG 235
121 ISEIAPSEIRGLNTLTPQSSGGMFLSYCVFQMSLSPSPDMRIMLVLAIPSLFFFG 180
236 TIFYPESPMLVSKGMAEKKVLOKLRGKDVSGELSLLEGEVGGDTISIEYIIGP 295
181 TIFYPESPMLVSKGMAEKKVLOKLRGKDVSGELSLLEGEVGGDTISIEYIIGP 240
296 ATEAADDLVTDGDKQITLYGPEEGOSWIAPSKGPIMLGSLVLSARHSGMNVQSPMD 355
241 ATEAADDLVTDGDKQITLYGPEEGOSWIAPSKGPIMLGSLVLSARHSGMNVQSPMD 300
356 DPTVTLFSGVHNNPQAGSGSRSTLFPNFGSMFVTDHANNQWDEBNLHRDDEBYASD 415
301 DPTVTLFSGVHNNPQAGSGSRSTLFPNFGSMFVTDHANNQWDEBNLHRDDEBYASD 360
416 GAGDYEENLSPLSRQATGABGKDIVHGHRSALSMRQTLLEGGGDVSTDIQGG 475
361 GAGDYEENLSPLSRQATGABGKDIVHGHRSALSMRQTLLEGGGDVSTDIQGG 420
476 WQAMKSEKENGKRGKFRVYLHOGVPGSRGSIYSLPGGDPFEGSEFVHAAL 535
421 WQAMKSEKENGKRGKFRVYLHOGVPGSRGSIYSLPGGDPFEGSEFVHAAL 480

Qy 536 VSQSALFSKGLAEPMSDAAVHPSEVAKGSRMKDLFEPGRALLVGVGIQIILQOFAG 595
 Db 481 VSQSALFSKGLAEPMSDAAVHPSEVAKGSRMKDLFEPGRALLVGVGIQIILQOFAG 540
 Qy 556 INGVYTPPQIIEQAGVAVILSKFGLSSASASISLITLMLPCIGPAMLMDSGR 655
 Db 541 INGVYTPPQIIEQAGVAVILSKFGLSSASASISLITLMLPCIGPAMLMDSGR 600
 Qy 656 FULLGTITPLIASLVLVNSNIDIGTLAHLSTVSIVVPCFVWGSPINILCAEI 715
 Db 601 FULLGTITPLIASLVLVNSNIDIGTLAHLSTVSIVVPCFVWGSPINILCAEI 660
 Qy 716 PPTVRGCLIAICAFETFGIIVVYSLPVMNAIGLAGVFSIYAVVCLISFVFLKVP 775
 Db 661 PPTVRGCLIAICAFETFGIIVVYSLPVMNAIGLAGVFSIYAVVCLISFVFLKVP 720
 Qy 776 ETKGMPLEVIITEFFAVGAKQAQAAKA 800
 Db 721 ETKGMPLEVIITEFFAVGAKQAQAAKA 745

RESULT 3
 US-10-051-902-2
 ; Sequence 2, Application US/10051902
 ; Publication No. US2002017846B1

GENERAL INFORMATION:
 APPLICANT: Allen, Steve
 APPLICANT: Hitz, Bill
 APPLICANT: Kinney, Tony
 APPLICANT: Tingey, Scott
 TITLE OF INVENTION: Plant Sugar Transport Proteins
 FILE REFERENCE: BB-1163
 CURRENT APPLICATION NUMBER: US/10/051,902
 PRIOR FILING DATE: 2002-01-17
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
 PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
 NUMBER OF SEQ ID NOS: 30
 SOFTWARE: Microsoft Office 97
 SEQ ID NO 2

Handwritten signature

LENGTH: 747
 TYPE: PRT
 ORGANISM: Zea mays
 FEATURE:
 NAME/KEY: UNSURE
 LOCATION: (129)
 NAME/KEY: UNSURE
 LOCATION: (133). (134)
 NAME/KEY: UNSURE
 LOCATION: (144)
 NAME/KEY: UNSURE
 LOCATION: (178)
 NAME/KEY: UNSURE
 LOCATION: (207)
 NAME/KEY: UNSURE
 LOCATION: (218)
 NAME/KEY: UNSURE
 LOCATION: (220)
 NAME/KEY: UNSURE
 LOCATION: (236)
 US-10-051-902-2

Query Match 86.3%; Score 3517; DB 14; Length 747;
 Beet Local Similarity 91.6%; Pred. No. 1.7e-305;
 Matches 684; Conservative 27; Mismatches 34; Indels 2; Gaps 2;

Qy 56 NSGAVLVAIVASIGNLQGMNATIAAVALYIKKEFQIQNEPTVEGLIVMSLIGATVIT 115
 Db 1 MGGAVLVAIVASIGNLQGMNATIAAVALYIKKEFQIQNEPTVEGLIVMSLIGATVIT 60
 Qy 116 TFSGLSDSISGRPMILISIIYFSGILMSPNVTYLLARFVDFGFGIATVILVPLY 175
 Db 61 TSPGRPADCGRRPMLVSAVALYFVSGIVMLWAPIYVILLARLIDGFGIGLATVILVPLY 120

Qy 176 ISEIAPSEIR-GLNLTLPQFSG-SGGMFLSYCMVFGNSLSPSPDMRIMLVIAIPSLPFF 233
 Db 121 ISEIAPSEIR-GLNLTLPQFSG-SGGMFLSYCMVFGNSLSPSPDMRIMLVIAIPSLPFF 180
 Qy 234 GLTFYLPESPRMIVSKRMAEAKKVLQKNGKDVSGELSLBLGELVGDTSIEEYI 293
 Db 181 GLTFYLPESPRMIVSKRMAEAKKVLQKNGKDVSGELSLBLGELVGDTSIEEYI 240
 Qy 294 GPATFAADLVTDGKEQITLYGPEEGQSWIARPSKGIIMGVSLASRRGSVNOVSVP 353
 Db 241 GPATFAADLVTDGKEQITLYGPEEGQSWIARPSKGIIMGVSLASRRGSVNOVSVP 300
 Qy 354 LMDPIVTLFGSVHNMPOAGSMRSTLPPNFGSMFVTDQAKRQOMDEMLHDDDEYA 413
 Db 301 LMDPIVTLFGSVHNMPOAGSMRSTLPPNFGSMFVTDQAKRQOMDEMLHDDDEYA 360
 Qy 414 SDGAGDYEDNLHSPILSRQATGAEKDIYVHGRGSLSMRQOTLLGEGDGVSTDIG 473
 Db 361 SDGAGDYEDNLHSPILSRQATGAEKDIYVHGRGSLSMRQOTLLGEGDGVSTDIG 420
 Qy 474 GGMOLAMKSEKEGNGRKEGFRKRVYIHOEGRVPSRRGSIIVSLPGGADVLEGEFVHAA 533
 Db 421 GGMOLAMKSEKEGNGRKEGFRKRVYIHOEGRVPSRRGSIIVSLPGGADVLEGEFVHAA 480
 Qy 534 ALVSQSALFSKGLAEPMSDAAVHPSEVAKGSRMKDLFEPGRALLVGVGIQIILQOF 593
 Db 481 ALVSQSALFSKGLAEPMSDAAVHPSEVAKGSRMKDLFEPGRALLVGVGIQIILQOF 540
 Qy 594 AGINGVLYTPPQIIEQAGVAVILSKFGLSSASASISLITLMLPCIGPAMLMDSG 653
 Db 541 AGINGVLYTPPQIIEQAGVAVILSKFGLSSASASISLITLMLPCIGPAMLMDSG 600
 Qy 654 RRFLLGTITPLIASLVLVNSNIDIGTLAHLSTVSIVVPCFVWGSPINILCAEI 713
 Db 601 RRFLLGTITPLIASLVLVNSNIDIGTLAHLSTVSIVVPCFVWGSPINILCAEI 660
 Qy 714 EIFPVRGCLIAICAFETFGIIVVYSLPVMNAIGLAGVFSIYAVVCLISFVFLK 773
 Db 661 EIFPVRGCLIAICAFETFGIIVVYSLPVMNAIGLAGVFSIYAVVCLISFVFLK 720
 Qy 774 VPEKMPLEVIITEFFAVGAKQAQAAKA 800
 Db 721 VPEKMPLEVIITEFFAVGAKQAQAAKA 745

Handwritten signature

RESULT 4
 US-10-051-909-2
 ; Sequence 2, Application US/10051909
 ; Publication No. US20020199217A1
 GENERAL INFORMATION:
 APPLICANT: Allen, Steve
 APPLICANT: Hellenjais, Tim
 APPLICANT: Hitz, Bill
 APPLICANT: Kinney, Tony
 APPLICANT: Tingey, Scott
 TITLE OF INVENTION: Plant Sugar Transport Proteins
 FILE REFERENCE: BB1163 US CIP
 CURRENT APPLICATION NUMBER: US/10/051,909
 PRIOR FILING DATE: 2002-01-17
 PRIOR APPLICATION NUMBER: 60/083,044
 PRIOR FILING DATE: April 24, 1998
 NUMBER OF SEQ ID NOS: 38
 SOFTWARE: Microsoft Office 97
 SEQ ID NO 2

LENGTH: 747
 TYPE: PRT
 ORGANISM: Zea mays
 FEATURE:
 NAME/KEY: UNSURE
 LOCATION: (129)
 NAME/KEY: UNSURE
 LOCATION: (133). (134)
 NAME/KEY: UNSURE

```

; LOCATION: (144)
; NAME/KEY: UNSURE
; LOCATION: (178)
; NAME/KEY: UNSURE
; LOCATION: (207)
; NAME/KEY: UNSURE
; LOCATION: (218)
; NAME/KEY: UNSURE
; LOCATION: (220)
; NAME/KEY: UNSURE
; LOCATION: (236)
; US-10-051-909-2

```

```

Query Match      86.3%; Score 3517; DB 14; Length 747;
Best Local Similarity 91.6%; Pred. No. 1.7e-305;
Matches 684; Conservative 27; Mismatches 34; Indels 2; Gaps 2;

```

```

QY 56 MSGAVLVAIVASIGNLQGMNATTAAGVLYIKKEPOLONEPTVEGLIVMSLIGATTIT 115
DB 1 MCGAVVVAIAISIGNLQGMNATTAAGVLYIKKEPOLONEPTVEGLIVMSLIGATTIT 60
QY 116 TFSGPLSDSISGRPMILISILYFSGIMLMSPNVYVLLARFVDFGFGIGLAATLVPLY 175
DB 61 TSPGRADOCVGRPMILVAVLYFVSGVLMAPYITILLARLIDGFGIGLAATLVPLY 120
QY 176 ISEIAPSEIR-GLANTLPQFSG-SCGMFLSYCMVFGMSLSPSDWIMLGVLAIPSLFFGL 233
DB 121 ISEIAPSEIR-GLANTLPQFSG-SCGMFLSYCMVFGMSLSPSDWIMLGVLAIPSLFFGL 180
QY 234 GLTITLPESSPRMVLVSKGMAEAKKYLQKLRGKDDVSGELSLLEGLFVGSDTSEIYII 293
DB 181 GLTITLPESSPRMVLVSKGMAEAKKYLQKLRGKDDVSGELSLLEGLFVGSDTSEIYII 240
QY 294 GPATEAADLVTDGKEQITLYGPEBGOSWIARPSKPIMLGSLVSLASRHGSMVNOQVP 353
DB 241 GPATEAADLVTDGKEQITLYGPEBGOSWIARPSKPIMLGSLVSLASRHGSMVNOQVP 300
QY 354 LMCPIVTLFGSVHEHNPQAGSMRSTLPFNFGSMFSVTDQAHAKNEQWDEENLHRDDEYA 413
DB 301 LMCPIVTLFGSVHEHNPQAGSMRSTLPFNFGSMFSVTDQAHAKNEQWDEENLHRDDEYA 360
QY 414 SDGAGDVEDNLHSPILSRQATGABGKDI-VHGHGGSALSMRQRTLLGEGGCVSSTDIG 473
DB 361 SDGAGDVEDNLHSPILSRQATGABGKDI-VHGHGGSALSMRQRTLLGEGGCVSSTDIG 420
QY 474 GGMOLAMKSEKGEKNGRKGKFRVYLLHQBEGVPSGRGSIYSLPGGDPVFGSEFVAAA 533
DB 421 GGMOLAMKSEKGEKNGRKGKFRVYLLHQBEGVPSGRGSIYSLPGGDPVFGSEFVAAA 480
QY 534 ALVSGSALFSGKLAEPMSDAAWHPSEVAAKGRMKDLPFGVRRALLVGVGIQLQOF 593
DB 481 ALVSGSALFSGKLAEPMSDAAWHPSEVAAKGRMKDLPFGVRRALLVGVGIQLQOF 540
QY 594 AGINGVLYTTPQILBOAGVAVILSKFGLSASASISLITLLMLPCIGFAMLMDSG 653
DB 541 AGINGVLYTTPQILBOAGVAVILSKFGLSASASISLITLLMLPCIGFAMLMDSG 600
QY 654 RFLILGTIPILIASVILVNSLIDGLTAHALSTVSIVYFCFVWGFPIPIILCA 713
DB 601 RFLILGTIPILIASVILVNSLIDGLTAHALSTVSIVYFCFVWGFPIPIILCA 660
QY 714 EFPTRVRGLCIAICAFETWIGDIIVTYSLPVMLNAGIAGVFSIYAVVCLISFVFLK 773
DB 661 EFPTRVRGLCIAICAFETWIGDIIVTYSLPVMLNAGIAGVFSIYAVVCLISFVFLK 720
QY 774 VPEIKGMPLEVTTEFPVAGKQAAKA 800
DB 721 VPEIKGMPLEVTTEFPVAGKQAAKA 747

```

RESULT 5
US-10-051-909-37
; Sequence 37, Application US/10051909

```

; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helenjaxis, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: B01163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 37
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Oryza sativa
; US-10-051-909-37

```

```

Query Match      69.5%; Score 2834; DB 14; Length 740;
Best Local Similarity 74.5%; Pred. No. 2.2e-244;
Matches 553; Conservative 75; Mismatches 108; Indels 6; Gaps 5;

```

```

QY 56 MSGAVLVAIVASIGNLQGMNATTAAGVLYIKKEPOLONEPTVEGLIVMSLIGATTIT 115
DB 1 MCGAVVVAIAISIGNLQGMNATTAAGVLYIKKEPOLONEPTVEGLIVMSLIGATTIT 60
QY 116 TFSGPLSDSISGRPMILISILYFSGIMLMSPNVYVLLARFVDFGFGIGLAATLVPLY 175
DB 61 TFSGAVASFRFRPMILIASAVLYFVSGVLMAPYITILLARLIDGFGIGLAATLVPLY 120
QY 176 ISEIAPSEIRLNTLPQFSGGGMFLSYCMVFGMSLSPSDWIMLGVLAIPSLFFGL 235
DB 121 ISEIAPSEIRLNTLPQFSGGGMFLSYCMVFGMSLSPSDWIMLGVLAIPSLFFGL 180
QY 236 TIFLIPESPRMVLVSKGMAEAKKYLQKLRGKDDVSGELSLLEGLFVGSDTSEIYII 295
DB 181 TIFLIPESPRMVLVSKGMAEAKKYLQKLRGKDDVSGELSLLEGLFVGSDTSEIYII 240
QY 296 ATBAADLVTDGDKQITLYGPEBGOSWIARPSKPIMLGSLVSLASRHGSMVNOQVP 355
DB 241 DDELABEGIAP-DPEKIKLYGPEBGOSWIARPSKPIMLGSLVSLASRHGSMVNOQVP 299
QY 356 DEIVTLFGSVHEHNPQAGSMRSTLPFNFGSMFSVTDQAHAKNEQWDEENLHRDDEYASD 415
DB 300 DEIVTLFGSVHEHNPQAGSMRSTLPFNFGSMFSVTDQAHAKNEQWDEENLHRDDEYASD 358
QY 416 GAGDYEENLHSPILSRQATGABGKDI-VHGHGGSALSMRQRTLLGEGGCVSSTDIG 474
DB 359 GAGDYEENLHSPILSRQATGABGKDI-VHGHGGSALSMRQRTLLGEGGCVSSTDIG 416
QY 475 GGMOLAMKSEKGEKNGRKGKFRVYLLHQBEGVPSGRGSIYSLPGGDPVFGSEFVAAA 534
DB 417 GGMOLAMKSEKGEKNGRKGKFRVYLLHQBEGVPSGRGSIYSLPGGDPVFGSEFVAAA 475
QY 535 LVSQSALFSGKLAEPMSDAAWHPSEVAAKGRMKDLPFGVRRALLVGVGIQLQOF 594
DB 476 LVSQSALFSGKLAEPMSDAAWHPSEVAAKGRMKDLPFGVRRALLVGVGIQLQOF 535
QY 595 GINGVLYTTPQILBOAGVAVILSKFGLSASASISLITLLMLPCIGFAMLMDSG 654
DB 536 GINGVLYTTPQILBOAGVAVILSKFGLSASASISLITLLMLPCIGFAMLMDSG 595
QY 655 RFLILGTIPILIASVILVNSLIDGLTAHALSTVSIVYFCFVWGFPIPIILCA 714
DB 596 RFLILGTIPILIASVILVNSLIDGLTAHALSTVSIVYFCFVWGFPIPIILCA 655
QY 715 IFTTRVRGLCIAICAFETWIGDIIVTYSLPVMLNAGIAGVFSIYAVVCLISFVFLK 774
DB 656 IFTTRVRGLCIAICAFETWIGDIIVTYSLPVMLNAGIAGVFSIYAVVCLISFVFLK 715

```


Qy 775 PETKGMPELVITEPFAVAKQA 796
 Db 716 PETKGMPELVITEPFAVAKQA 737

RESULT 6

US-10-051-902-8
 ; Sequence 8, Application US/10051902
 ; Publication No. US20020178468A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Hite, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB-1163
 ; CURRENT APPLICATION NUMBER: US/10/051,902
 ; CURRENT FILING DATE: 2002-01-17
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 8
 ; LENGTH: 737
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; US-10-051-902-8

Query Match 65.6%; Score 2674; DB 14; Length 737;
 Best Local Similarity 69.9%; Pred. No. 4.7e-230;
 Matches 522; Conservative 92; Mismatches 115; Indels 18; Gaps 8;

Qy 56 MSGAVLVAIVASIGNLLQGMNATIAAVALYIKKEFOLQNEPTVEGLIVSMISLIGATVIT 115
 Db 1 MKGAVALVAIAASIGNFLQGMNATIAAGANGYIKOLALGT--TWERLVVGMISLIGATVIT 58
 Qy 116 TFSGLPSISGRRLPMLISLIFPSGLIMLSPNVYLLARFVDFGIGLATVLPVLY 175
 Db 59 TCSGIAIADWLRPRMTISSVLYFLGLVLMSPNVYLLCLARLDGFGIGLATVLPVLY 118
 Qy 176 ISEIAPSEIRGLNLTLPQFSGSGMFLSYCMVFGMSLSPSPDWRIMLGVALIPSLFFPGL 235
 Db 119 ISEIAPSEIRGLNLTLPQFSGSGMFLSYCMVFGMSLSPSPDWRIMLGVALIPSLFFPGL 178
 Qy 236 TIFPLPESPRMLVSKGRMAEAKVYLQRLRGKDVSGELSLLEGLLEVGDTSIEEYITGP 295
 Db 179 TIFPLPESPRMLVSKGRMAEAKVYLQRLRGKDVSGELSLLEGLLEVGDTSIEEYITGP 238
 Qy 296 ATEAADDLVTDGDKQITLYGPEEGSWIAPRSKGPIMLGSVLSASRHSVMNQSVPLM 355
 Db 239 ADVADGDHGHATEKDKIRLYSGQAGLSWLSKPVYTGSSIG---LASHHGSIINOSMPLM 294
 Qy 356 DPIVTLFGSVHNNPQAG--GSMRSTLPNFGSMFVTDQAKNQEWDENLHRDDEEYA 413
 Db 295 DPLVTLFGSIHEKLPETGARGSMRSTLPNFGSMFVTDQAKNQEWDENLHRDDEEYA 354
 Qy 414 SDGAGGYEDNLHSLPSLRQATGAEKDIYHNGRGSAL--SMRQTLLEG--GDGVSTG 471
 Db 355 SDATRGDSDNLHSLPSLRQATGAEKDIYHNGRGSAL--SMRQTLLEG--GDGVSTG 413
 Qy 472 IGGGQOLAMKMSSEKGENGRKGGFKRYVYLHQBEGVPSRGSIVSLPGGADVFESEFVH 531
 Db 414 IGGGQOLAMKMTDK--GEGDKQGGFKRYVYLHQBEGVPSRGSIVSLPGGADVFESEFVH 466
 Qy 532 AAALVSGALPFSKGLAPRMSDAMVHPSEVAAGSKWKDLFEFGVRALLVGVGIQILQ 591
 Db 467 AAALVSGALPFSKGLAPRMSDAMVHPSEVAAGSKWKDLFEFGVRALLVGVGIQILQ 526
 Qy 592 QPAGINGVLYTTPOLLEQAGAVIISKFGLSASASILISLTTLLMPCIGFAMLLNDL 651
 Db 527 QPAGINGVLYTTPOLLEQAGAVIISKFGLSASASILISLTTLLMPCIGFAMLLNDL 586
 Qy 652 SGRRLGLTTPILIASVILVSNLIDGLTAHALLSIVSVIYVFCFVWGFGPIPNIL 711

Db 587 SGRRLGLTTPILIASVILVSNLIDGLTAHALLSIVSVIYVFCFVWGFGPIPNIL 646
 Qy 712 CAEIFFTRVGLCAICAFPMWIGDITVYSLPMLNATIGLAVFSLIYAVVCLISFVFE 771
 Db 647 CSEIFFTRVGLCAICAFPMWIGDITVYSLPMLNATIGLAVFSLIYAVVCLISFVFE 706
 Qy 772 LKVPETKGMPELVITEPFAVAKQA 798
 Db 707 LKVPETKGMPELVITEPFAVAKQA 733

RESULT 7

US-10-051-909-8
 ; Sequence 8, Application US/10051909
 ; Publication No. US20020199217A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Hite, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB1163 US CIP
 ; CURRENT APPLICATION NUMBER: US/10/051,909
 ; CURRENT FILING DATE: 2002-01-17
 ; PRIOR APPLICATION NUMBER: 60/083,044
 ; PRIOR FILING DATE: April 24, 1998
 ; NUMBER OF SEQ ID NOS: 38
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 8
 ; LENGTH: 737
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; US-10-051-909-8

Query Match 65.6%; Score 2674; DB 14; Length 737;
 Best Local Similarity 69.9%; Pred. No. 4.7e-230;
 Matches 522; Conservative 92; Mismatches 115; Indels 18; Gaps 8;

Qy 56 MSGAVLVAIVASIGNLLQGMNATIAAVALYIKKEFOLQNEPTVEGLIVSMISLIGATVIT 115
 Db 1 MKGAVALVAIAASIGNFLQGMNATIAAGANGYIKOLALGT--TWERLVVGMISLIGATVIT 58
 Qy 116 TFSGLPSISGRRLPMLISLIFPSGLIMLSPNVYLLARFVDFGIGLATVLPVLY 175
 Db 59 TCSGIAIADWLRPRMTISSVLYFLGLVLMSPNVYLLCLARLDGFGIGLATVLPVLY 118
 Qy 176 ISEIAPSEIRGLNLTLPQFSGSGMFLSYCMVFGMSLSPSPDWRIMLGVALIPSLFFPGL 235
 Db 119 ISEIAPSEIRGLNLTLPQFSGSGMFLSYCMVFGMSLSPSPDWRIMLGVALIPSLFFPGL 178
 Qy 236 TIFPLPESPRMLVSKGRMAEAKVYLQRLRGKDVSGELSLLEGLLEVGDTSIEEYITGP 295
 Db 179 TIFPLPESPRMLVSKGRMAEAKVYLQRLRGKDVSGELSLLEGLLEVGDTSIEEYITGP 238
 Qy 296 ATEAADDLVTDGDKQITLYGPEEGSWIAPRSKGPIMLGSVLSASRHSVMNQSVPLM 355
 Db 239 ADVADGDHGHATEKDKIRLYSGQAGLSWLSKPVYTGSSIG---LASHHGSIINOSMPLM 294
 Qy 356 DPIVTLFGSVHNNPQAG--GSMRSTLPNFGSMFVTDQAKNQEWDENLHRDDEEYA 413
 Db 295 DPLVTLFGSIHEKLPETGARGSMRSTLPNFGSMFVTDQAKNQEWDENLHRDDEEYA 354
 Qy 414 SDGAGGYEDNLHSLPSLRQATGAEKDIYHNGRGSAL--SMRQTLLEG--GDGVSTG 471
 Db 355 SDATRGDSDNLHSLPSLRQATGAEKDIYHNGRGSAL--SMRQTLLEG--GDGVSTG 413
 Qy 472 IGGGQOLAMKMSSEKGENGRKGGFKRYVYLHQBEGVPSRGSIVSLPGGADVFESEFVH 531
 Db 414 IGGGQOLAMKMTDK--GEGDKQGGFKRYVYLHQBEGVPSRGSIVSLPGGADVFESEFVH 466
 Qy 532 AAALVSGALPFSKGLAPRMSDAMVHPSEVAAGSKWKDLFEFGVRALLVGVGIQILQ 591

Db 467 AALATVSRPLYSKELIDGHPVGRPAWHPSETASKGSMALLPEPVKALVVGVIQIIQ 526
Qy 552 QFAGINGVLYTPPQILEQAVVILSKFGLSASASILISLTLLMPCIGPAMLM 651
Db 527 QPSGNGVLYTPPQILEQAVVILSKFGLSASASILISLTLLMPCIGPAMLM 586
Qy 652 SGRRLGLTIPILIASVILVNSNIDGLTAHALSTVSIYVFCFPMGFGPIL 711
Db 587 SGRRLGLTIPILIASVILVNSNIDGLTAHALSTVSIYVFCFPMGFGPIL 646
Qy 712 CAEIPTRVGLCIAICAFTEWIGDIIVTSLPVMNAIGLAVFSIYAVCLISFV 771
Db 647 CSEIPTRVGLCIAICAFTEWIGDIIVTSLPVMNAIGLAVFSIYAVCLISFV 706
Qy 772 LKVPETKGMPLVITTEFFAVGAKQAAA 798
Db 707 LKVPETKGMPLVITTEFFAVGAKQAAA 733

RESULT 8

US-10-051-902-29
; Sequence 29, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 29
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-051-902-29

Query Match 63.4%; Score 2584; DB 14; Length 729;
Best Local Similarity 68.8%; Pred. No. 5.2e-222;
Matches 512; Conservative 89; Mismatches 123; Indels 20; Gaps 8;

Qy 56 MSGAVLVAIVASISGNLQGMNATIAAVALYIKKEFQIQNEPTVEGLIYMSLIGATV 115
Db 1 MSGAVLVAIVASISGNLQGMNATIAAVALYIKKEFQIQNEPTVEGLIYMSLIGATV 60
Qy 116 TFSGLSDSIGRRPMLISIIYFSGILMSPNVYVLLARFVDFGIGLAVTLVPL 175
Db 61 TCSGCVADWLCRRPMLISIIYFSGILMSPNVYVLLARFVDFGIGLAVTLVPL 120
Qy 176 ISEIAPBEIRGLANTLPQFSGGMPFLSYCMVFCGMSLSPSPDMRIMLGVALIPSLFF 235
Db 121 ISEIAPBEIRGLANTLPQFSGGMPFLSYCMVFCGMSLSPSPDMRIMLGVALIPSLFF 180
Qy 236 TIFLPSRPMVLVSKGMALAKRYLQRLRGKDVSGELSLLEGLVEGDTSIIEYIIGP 295
Db 181 TIFLPSRPMVLVSKGMALAKRYLQRLRGKDVSGELSLLEGLVEGDTSIIEYIIGP 240
Qy 296 ATEAADLVTDGDEQITLYGPEEGOSWIARPSKPIMLGVSILASRHGS-MVNQSVPL 354
Db 241 ATEAADLVTDGDEQITLYGPEEGOSWIARPSKPIMLGVSILASRHGS-MVNQSVPL 296
Qy 355 MDPIYTLGSGVHNNPQAGGSMRSLTFPRGSMESVTQOAHKEQWDEENLHRDEEYAS 414
Db 297 IDPLVTLGSGVHNNPQAGGSMRSLTFPRGSMESVTQOAHKEQWDEENLHRDEEYAS 355
Qy 415 DGAGDVEDNHLHSPLSRQATGAEGKDIYHGHRSALSMRQTLIGGSDGVSSTDIGG 474

Db 356 D-HGDDSDDLHSLPISIKQTTSME-KOMPHTAGTISTFRHSGVQQAQSGAGSMGICG 413
Qy 475 GWOIAMKSEKEGENGRKGGFKRYLHOGGVPSGRGSIYSLPGGDFEFGSEFVHAAA 534
Db 414 GWOIAMKSEKEGENGRKGGFKRYLHOGGVPSGRGSIYSLPGGDFEFGSEFVHAAA 462
Qy 535 LVSOALFESKGLAPRMSDAMVHPSEVAAGSMKDLFEEGVARALLVGVGIILOOFA 594
Db 463 LVSOALFESKGLAPRMSDAMVHPSEVAAGSMKDLFEEGVARALLVGVGIILOOFA 521
Qy 595 GINGVLYTPPQILEQAVVILSKFGLSASASILISLTLLMPCIGPAMLM 654
Db 527 GINGVLYTPPQILEQAVVILSKFGLSASASILISLTLLMPCIGPAMLM 581
Qy 655 RPLIGTIPILIASVILVNSNIDGLTAHALSTVSIYVFCFPMGFGPIL 714
Db 582 RPLIGTIPILIASVILVNSNIDGLTAHALSTVSIYVFCFPMGFGPIL 641
Qy 715 IEPTRVGLCIAICAFTEWIGDIIVTSLPVMNAIGLAVFSIYAVCLISFV 774
Db 642 IEPTRVGLCIAICAFTEWIGDIIVTSLPVMNAIGLAVFSIYAVCLISFV 701
Qy 775 PETKGMPLVITTEFFAVGAKQAAA 798
Db 702 PETKGMPLVITTEFFAVGAKQAAA 735

RESULT 9

US-10-051-909-29
; Sequence 29, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helencjatis, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 29
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-051-909-29

Query Match 63.4%; Score 2584; DB 14; Length 729;
Best Local Similarity 68.8%; Pred. No. 5.2e-222;
Matches 512; Conservative 89; Mismatches 123; Indels 20; Gaps 8;

Qy 56 MSGAVLVAIVASISGNLQGMNATIAAVALYIKKEFQIQNEPTVEGLIYMSLIGATV 115
Db 1 MSGAVLVAIVASISGNLQGMNATIAAVALYIKKEFQIQNEPTVEGLIYMSLIGATV 60
Qy 116 TFSGLSDSIGRRPMLISIIYFSGILMSPNVYVLLARFVDFGIGLAVTLVPL 175
Db 61 TCSGCVADWLCRRPMLISIIYFSGILMSPNVYVLLARFVDFGIGLAVTLVPL 120
Qy 176 ISEIAPBEIRGLANTLPQFSGGMPFLSYCMVFCGMSLSPSPDMRIMLGVALIPSLFF 235
Db 121 ISEIAPBEIRGLANTLPQFSGGMPFLSYCMVFCGMSLSPSPDMRIMLGVALIPSLFF 180
Qy 236 TIFLPSRPMVLVSKGMALAKRYLQRLRGKDVSGELSLLEGLVEGDTSIIEYIIGP 295
Db 181 TIFLPSRPMVLVSKGMALAKRYLQRLRGKDVSGELSLLEGLVEGDTSIIEYIIGP 240
Qy 296 ATEAADLVTDGDEQITLYGPEEGOSWIARPSKPIMLGVSILASRHGS-MVNQSVPL 354

Db 241 ADEVTDDHDIADVQDKIYLGAEBGLSWVARPVK-----GSTMSVLSRHSGTMSRRQSL 296
Qy 355 MDPIVTLFGSVHNNPQAGGSMRSTLPPNFGSMFVTDQHAKEQWDEENLHRDEEYAS 414
Db 297 IDPLVTLFGSVHNNPDT-GSMRSALFPFGSMFVGNQPRHEDWDEENLHVBEGEDYPS 355
Qy 415 DGAGDYEDNLHSPILSRQATGABGKDIVHGHRSALSMRQTLLEGSGDVSSTDIG 474
Db 356 D-HGDDSDDLHSPILSRQTTSM-KDMPHTAHGLSTFRHSGYQAGQBGASGMGCG 413
Qy 475 GQWLAMKSKSEKENGKRGKRVYLVHOGVPGSRGSIYSLPGGDDVPEGSEFVHAAA 534
Db 414 GQVVMKMTTERDESGQKE-----EGFPGSRGSIYSLPGGDDTGE-ADFFVQASA 462
Qy 535 LVSQALSKGLAEPRMDAAMVHPSEVAAGSRKMDLPEPGVRALLVGVGIOLLOPA 594
Db 463 LVSQALSKGLKHTTIGPAWHPSE-TTKGSIHMDLHPGVKALVGVGLQILOPS 521
Qy 595 GINGVLYTTPQLEQAGVAVILSKFGLSASASILISLTLLMLPCIGFAMLLMDLSGR 654
Db 522 GINGVLYTTPQLEQAGVAVILSKFGLSASASILISLTLLMLPALAVAMRLMDLSGR 581
Qy 655 RPLIGTIPILIASVILVSNLIDGLTAAHALSTVSIVYVCCFVWGFGPIENILCAE 714
Db 582 RPLITITPILIASVILVSNLIDGLTAAHALSTVSIVYVCCFVWGFGPAPILGSE 641
Qy 715 IFFRVRGICIAICFTFMIGDIIVTYSIPVMLNAGLAVSIAVVCILSFVFPFLK 774
Db 642 IFFRVRGICIAICFTFMIGDIIVTYSIPVMLNAGLAVSIAVVCILSWFVFLK 701
Qy 775 PETKGMPLVITEFPFVAGAKQAAA 798
Db 702 PETKGMPLVITEFPFVAGAKQAAA 725

RESULT 10
US-10-051-902-10
; Sequence 10, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tiney, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Glycine max
US-10-051-902-10

Query Match 41.0%; Score 1670.5; DB 14; Length 486;
Best Local Similarity 68.1%; Pred. No. 1.6e-140;
Matches 333; Conservative 57; Mismatches 84; Indels 15; Gaps 8;

Qy 308 DKEQITLYGPEEGSWIARPSKGPIMLSVLSASRHGMNOSVPLMDPIYTLFGSVH 367
Db 5 EKDQIKLYGPEEGSWIARPSKGPIMLSVLSASRHGMNOSVPLMDPIYTLFGSVH 59
Qy 368 NMPQAGSMRSTLPPNFGSMFVTDQHAKEQWDEENLHRDEEYASDAGDYEDNLH 427
Db 60 KLPBEG-----STLFPFGSMFVGNQPRHEDWDEESLARBGDDYSD--AGDSNDLQS 113
Qy 428 PLSRQATGABGKDIVHGHRSALSMRQTL-GEAGDVSTDIIGGQWLAMKSKSE 486
Db 114 PLSRQATSLD-KDIPPAHSNLA-SMRQSLHGNSGPTSGTIGGQWLAMKSKSE 171

Qy 487 GENGKKEGKFRVYLVHOGVPGSRGSIYSLPGGDDVPEGSEFVHAAALVQSALFSKGL 546
Db 172 GPDGKKEGKFRVYLVHOGVPGSRGSIYSLPGGDDVPEGSEFVHAAALVQSALFSKGL 230
Qy 547 AEPMSDAAVHPSEVAAGSRKMDLPEPGVRALLVGVGIOLLOPAGINGVLYTPOI 606
Db 211 MRQRPVGPAMHPEBTIAKPSWMDLPEPGVKAHALVGVGIOLLOPAGINGVLYTPOI 290
Qy 607 LEQAGVAVILSKFGLSASASILISLTLLMLPCIGFAMLLMDLSGRFLLGTIPLI 666
Db 291 LEQAGVAVILSKFGLSASASILISLTLLMLPCIGFAMLLMDLSGRFLLGTIPLI 350
Qy 667 ASVILVSNLIDGLTAAHALSTVSIVYVCCFVWGFGPIENILCAIFPTRRGICIA 726
Db 351 ASVILVSNLIDGLTAAHALSTVSIVYVCCFVWGFGPIENILCAIFPTRRGICIA 410
Qy 727 ICAFTFMIGDIIVTYSIPVMLNAGLAVSIAVVCILSFVFPFLKVPETKGMPLVIT 786
Db 411 ICAFTFMIGDIIVTYSIPVMLNAGLAVSIAVVCILSFVFPFLKVPETKGMPLVIT 470
Qy 787 EFPVAGAKQ 795
Db 471 EFPVAGAKQ 479

RESULT 11
US-10-051-909-10
; Sequence 10, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helencjars, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tiney, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Glycine max
US-10-051-909-10

Query Match 41.0%; Score 1670.5; DB 14; Length 486;
Best Local Similarity 68.1%; Pred. No. 1.6e-140;
Matches 333; Conservative 57; Mismatches 84; Indels 15; Gaps 8;

Qy 308 DKEQITLYGPEEGSWIARPSKGPIMLSVLSASRHGMNOSVPLMDPIYTLFGSVH 367
Db 5 EKDQIKLYGPEEGSWIARPSKGPIMLSVLSASRHGMNOSVPLMDPIYTLFGSVH 59
Qy 368 NMPQAGSMRSTLPPNFGSMFVTDQHAKEQWDEENLHRDEEYASDAGDYEDNLH 427
Db 60 KLPBEG-----STLFPFGSMFVGNQPRHEDWDEESLARBGDDYSD--AGDSNDLQS 113
Qy 428 PLSRQATGABGKDIVHGHRSALSMRQTL-GEAGDVSTDIIGGQWLAMKSKSE 486
Db 114 PLSRQATSLD-KDIPPAHSNLA-SMRQSLHGNSGPTSGTIGGQWLAMKSKSE 171
Qy 487 GENGKKEGKFRVYLVHOGVPGSRGSIYSLPGGDDVPEGSEFVHAAALVQSALFSKGL 546
Db 172 GPDGKKEGKFRVYLVHOGVPGSRGSIYSLPGGDDVPEGSEFVHAAALVQSALFSKGL 230
Qy 547 AEPMSDAAVHPSEVAAGSRKMDLPEPGVRALLVGVGIOLLOPAGINGVLYTPOI 606
Db 211 MRQRPVGPAMHPEBTIAKPSWMDLPEPGVKAHALVGVGIOLLOPAGINGVLYTPOI 290

QY 607 LEQAGVAVILSKFGSLSSASASIIISLTLLMPCIGFAMLMDSGRFFLLGTPIII 666
DB 291 LEQAGVGLYSLIGLSTSSSSFLISAVTLLMPCIAIMRLMDISGRRTLLSTPIVI 350
QY 667 ASLVILVSNLIDLGTLAAALSTVAVIYFPCFVWGSPGPIINILCAEFPTVRGLCIA 726
DB 351 ALMLILVLSLDVDSSTANASISTTSVIYFCFVWGSPGPIINILCAEFPTVRGLCIA 410
QY 727 ICAFTFWIGDIIVTYSLPVMAIAGLAVFSIYAVVCLISFVFLKVBETKMPLEVT 786
DB 411 ICAFTFWICDIIVTYSLPVMAIAGLAVFSIYAVVCLISFVFLKVBETKMPLEVT 470
QY 787 EFPAVGAQK 795
DB 471 EFPAVGAQK 479

RESULT 12

US-10-051-902-14
; Sequence 14, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-051-902-14

Query Match 34.9%; Score 1424; DB 14; Length 345;
Best Local Similarity 82.9%; Pred. No. 1.1e-118;
Matches 281; Conservative 24; Mismatches 32; Indels 2; Gaps 2;

QY 462 EGGDGVSTIDIGGQWLAKWSEKENGKRGKRYVYLHOGVPGSRGSIYSLPGG 521
DB 4 EGGBAVSTIGIGGQWLAKWSEKENGKRGKRYVYLHOGVPGSRGSIYSLPGG 63
QY 522 DVEF-GSEFVAALVSGALFSGKLAEPFM-SDAAMVHPSEVAAKGRWDLFEPGVR 579
DB 64 DATGGSGFTIHAALVSHSALYSKDLMBERMAAGPMIHPLEAPKGSIMKDLFEPGVR 123
QY 580 ALVGVGIILOQFAGINVLVYTPQILEQAGVAVILSKFGSLSSASIIISLTLLM 639
DB 124 ALFVGVGIILOQFAGINVLVYTPQILEQAGVAVILSKFGSLSSASIIISLTLLM 183
QY 640 PCIGFAMLMDSGRFFLLGTPIILASLVILVSNLIDLGTLAAALSTVAVIYFCC 699
DB 184 PSIGVAMRLMDISGRFFLLGTPIILASLVILVSNLIDLGTLAAALSTVAVIYFCC 243
QY 700 FVMGSPPIINILCAEFPTVRGLCIAFTFWIGDIIVTYSLPVMAIAGLAVFSIY 759
DB 244 FVMGSPPIINILCAEFPTVRGLCIAFTFWIGDIIVTYSLPVMAIAGLAVFSIY 303
QY 760 AVVCLISFVFLKVBETKMPLEVTTEFAVGAQKAAA 798
DB 304 AVVCCIAFVFLKVBETKMPLEVTTEFAVGAQKAAA 342

RESULT 13
US-10-051-909-14
; Sequence 14, Application US/10051909

Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Heltentaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-051-909-14

Query Match 34.9%; Score 1424; DB 14; Length 345;
Best Local Similarity 82.9%; Pred. No. 1.1e-118;
Matches 281; Conservative 24; Mismatches 32; Indels 2; Gaps 2;

QY 462 EGGDGVSTIDIGGQWLAKWSEKENGKRGKRYVYLHOGVPGSRGSIYSLPGG 521
DB 4 EGGBAVSTIGIGGQWLAKWSEKENGKRGKRYVYLHOGVPGSRGSIYSLPGG 63
QY 522 DVEF-GSEFVAALVSGALFSGKLAEPFM-SDAAMVHPSEVAAKGRWDLFEPGVR 579
DB 64 DATGGSGFTIHAALVSHSALYSKDLMBERMAAGPMIHPLEAPKGSIMKDLFEPGVR 123
QY 580 ALVGVGIILOQFAGINVLVYTPQILEQAGVAVILSKFGSLSSASIIISLTLLM 639
DB 124 ALFVGVGIILOQFAGINVLVYTPQILEQAGVAVILSKFGSLSSASIIISLTLLM 183
QY 640 PCIGFAMLMDSGRFFLLGTPIILASLVILVSNLIDLGTLAAALSTVAVIYFCC 699
DB 184 PSIGVAMRLMDISGRFFLLGTPIILASLVILVSNLIDLGTLAAALSTVAVIYFCC 243
QY 700 FVMGSPPIINILCAEFPTVRGLCIAFTFWIGDIIVTYSLPVMAIAGLAVFSIY 759
DB 244 FVMGSPPIINILCAEFPTVRGLCIAFTFWIGDIIVTYSLPVMAIAGLAVFSIY 303
QY 760 AVVCLISFVFLKVBETKMPLEVTTEFAVGAQKAAA 798
DB 304 AVVCCIAFVFLKVBETKMPLEVTTEFAVGAQKAAA 342

RESULT 14

US-10-051-902-16
; Sequence 16, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-051-902-16

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 5, 2004, 19:05:50 ; Search time 4174 Seconds

(without alignments)
4658.261 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGSWLAVQPPFPDIDRR.....PLEVITFEFVAGKAAAKA 800

Scoring table:
BLOSUM62
Xgapop 10.0, Ygapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlh
-O=/cgm2_1/USPTO.spool/US10051909/runat_05012004_094742_22317/app_query.fasta_1.967
-DB=EST -QFMT=fastap -SUFFIX=est -MINMATCH=0.1 -LOEPL=0 -LOEEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR SCORE=500 -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051909@cgm2_1 12937 @runat_05012004_094742_22317 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fut:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3027	74.3	2180	11 AY105508	AY105508 Zea mays
2	1399	34.3	911	29 CC423708	CC423708 PUFOR49TB
3	1382.5	33.9	845	29 BZ723844	BZ723844 PUCF160TD
4	1279	31.4	795	29 BZ989776	BZ989776 PUCD907TD
5	1101	27.0	829	13 BQ752737	BQ752737 WBE4118_E
6	990.5	24.3	745	14 CB635485	CB635485 OST1EB150
7	981	24.1	594	9 A1861088	A1861088 603011H02
8	957	23.5	817	29 BZ819234	BZ819234 PUGDY07TB
9	930	22.8	859	12 BM817360	BM817360 HCl06B05
10	928	22.8	749	13 BU040051	BU040051 PP_LB8000
11	924	22.7	688	14 CA502649	CA502649 WBE4338_A
12	921.5	22.6	700	13 BQ870435	BQ870435 QGD9A19_Y
13	921.5	22.6	703	14 CA502490	CA502490 WBE4048_B
14	916	22.5	847	14 CB981615	CB981615 CAB70005
15	909	22.3	572	12 BM325261	BM325261 PIC1_42_E
16	908	22.3	762	12 B1310999	B1310999 EST531274
17	902.5	22.1	710	13 BQ870328	BQ870328 QGD8L23_Y
18	895.5	22.0	738	14 CB965602	CB965602 NL23_F11
19	894	21.9	645	14 CA030046	CA030046 HK05020r
20	892.5	21.9	727	12 BJ471780	BJ471780 BJ471780
21	891	21.9	682	13 BU007863	BU007863 QCH5106_Y
22	889	21.8	671	9 AV909064	AV909064 AV909064
23	888.5	21.8	776	14 CB671921	CB671921 OSUNB05L
24	877.5	21.5	739	12 B1920655	B1920655 EST540578
25	876	21.5	656	13 BU925675	BU925675 MCS015A08
26	875	21.5	688	13 BU004365	BU004365 OCG4A09_Y
27	874.5	21.5	805	14 CB673272	CB673272 OSUNB07L
28	865	21.2	653	12 BU450590	BU450590 BU450590
29	855	21.0	682	13 BU040035	BU040035 PP_LB8000
30	852	20.9	720	12 B1310710	B1310710 EST531246
31	850	20.9	819	10 BG584543	BG584543 EST5486304
32	849.5	20.8	706	12 BJ472944	BJ472944 BJ472944
33	849	20.8	624	14 CB213154	CB213154 OMT03434
34	847.5	20.8	787	14 CB634320	CB634320 OST1EB13M
35	840	20.6	617	13 BU991448	BU991448 HD06P09r
36	838	20.6	679	13 BU040591	BU040591 PP_LB8000
37	837.5	20.6	620	13 BU989641	BU989641 HP22E20r
38	836	20.5	609	12 B1309579	B1309579 EST530989
39	834.5	20.5	675	13 BU989051	BU989051 HP19J23r
40	833	20.4	725	13 BQ858206	BQ858206 QCB9K03_Y
41	831	20.4	627	14 CB816559	CB816559 3529_1_80
42	830.5	20.4	673	12 BJ449118	BJ449118 BJ449118
43	827	20.3	645	13 BQ240484	BQ240484 TAE05016G
44	825.5	20.3	751	13 BO993492	BO993492 OGR3f04_Y
45	824.5	20.2	699	14 CB966427	CB966427 NL37_F05

ALIGNMENTS

RESULT 1
AY105508
LOCUS AY105508 2180 bp mRNA linear HTC 16-OCT-2002
DEFINITION Zea mays PC0114533 mRNA sequence.
ACCESSION AY105508
VERSION AY105508.1 GI:21208586
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 2180)

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
SOURCE

Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whiteitt,M.S.,
Arthur,L.M., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 2180)
Coe,E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSU, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schubler, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
Location/Qualifiers
1. 2180
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="maizedb:63667"
/db_xref="taxon:4577"
/clone_id="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

BASE COUNT 533 a 447 c 555 g 645 t -
ORIGIN

Alignment Scores:
Pred. No.: 5.93e-274 Length: 2180
Score: 3027.00 Matches: 591
Percent Similarity: 99.66% Conservative: 2
Best Local Similarity: 99.33% Mismatches: 2
Query Match: 74.28% Indels: 0
DB: Gaps: 0

US-10-051-909-32 (1-800) x AY105508 (1-2180)

QY 205 MetValPheGlyMetSerLeuSerProSerProAspTrpArgIleMetLeuValIleu 225
DB 50 ATGGTGTGGAGATGCTGTCGCGCATCCAGCAATGAGAAATGCTGGTGCTC 109
QY 226 AlAlieProSerLeuPhePhePheGlyLeuThrIlePheTyrlleuProGlySerProArg 245
DB 110 GCGATACCTTCATGTTCTTCTTGGTTGACAAATATTTATCTTCTGATCTCCAGA 169
QY 246 TrpLeuValSerIleGlyArgMetAlaGluAlaIleValIleValIleValIleVal 265
DB 170 TGGCTCGTTGCAAGGTCGAGTGCAGAGCAAAAAGGTTGCAAAAGTTACCGTGC 229
QY 266 LysAspAspValSerIleGlyIleuSerLeuLeuLeuIleGlyLeuGlyValIleGlyAsp 285
DB 230 AAACACCATGCTCAGGTAATGTCCTTCTTCTCAAGGGTTGGAGGTTGAGGAGAC 289
QY 286 ThrSerIleGlyIleGlyIleIleGlyProAlaThrGluAlaIleAspAspLeuValThr 305
DB 290 ACTTCCATTGAGAGTATCATCATGAGCTGAGCTCCAGGACGCGATGATCTTGTACT 349
QY 306 AspGlyAspIleGlyIleIleThrLeuTyrlleProGlyIleGlyIleSerTrpIleAla 325
DB 350 GACGATATAGGAACAATACACTTATGGGCTGAGAGAGCCGATGATGATGCT 409
QY 326 ArgProSerIleGlyProIleMetLeuGlySerValIleSerLeuAlaSerArgHisGly 345
DB 410 CGACCTTCTAAGGAGCCATCATCTTGGAGAGTGTCTTCTTGTGATCTCGCATGGG 469
QY 346 SerMetValAsnIleSerValProLeuMetAspProIleValThrLeuPheGlySerVal 365

DB 470 AGCATGGTGAACCAAGATGATCCCTTATGATCGATGTCATCTTTGTAGTGTGC 529
QY 366 HisGluAsnMetProGluAlaGlyIleSerMetArgSerThrLeuPheProAspPheGly 385
DB 530 CATGAGAAATATGCTCCCAAGCTGAGAGAAATGAGGACATGTTCTCCAACTTTGGA 589
QY 386 SerMetPheSerValThrAspGluHisAlaIleValSerLeuIleTrpAspGluValAsnLeu 405
DB 590 AGTATGTTCACTGTCACAGATCAGATCCAAAATAGCATGCTGGATGAAAGAAATCTT 649
QY 406 HisArgAspAspGlyIleGlyIleAlaSerAspGlyAlaGlyIleAspTyrlleValAspAsnLeu 425
DB 650 CATAGGATGACGAGAGATGACATGATGCTGAGAGAGTGACTATGAGCAATCTC 709
QY 426 HisSerProLeuLeuSerArgGluAlaThrGlyAlaGlyIleValIleHis 445
DB 710 CATACCCCATTTGCTGTCCAGGACAGCAACAGTCCGAGGAGGAGCAATTTGTCACCAT 769
QY 446 GlyHisArgGlySerAlaLeuSerMetArgArgIleThrLeuLeuGlyIleGlyIleAsp 465
DB 770 GGTACCGTGAAGTGTCTTTAGCATGAGAAAGCAAGCCCTTTAGGGAGGCTGAGAT 829
QY 466 GlyValSerSerThrAspIleGlyIleGlyIleProIleuAlaTrpIleTrpSerGlyIle 485
DB 830 GGTGTGACAGACATGATATCGGTGGGAGTGCAGCTTCTTGGAAATGTCAGAGAG 889
QY 486 GlyIleGlyIleAsnGlyIleValGlyIleGlyIleValIleValIleValIleValIle 505
DB 890 GAAGGTGAGAAATGTGAGAAAGAGTGTCTTCAAAAGATCTTGCACCAAGAGGA 949
QY 506 ValProGlySerArgArgGlySerIleValSerIleProGlyIleGlyIleAspValPheGly 525
DB 950 GTTCTGCTCAAGAGGAGGCTCAATTTTTCATCTCCGGTGTGGGATGTTCTTGAG 1009
QY 526 GlySerGluPheValHisAlaAlaLeuValSerIleSerAlaLeuPheSerIleGly 545
DB 1010 GGTAGTGAATTTGTCATGCTGCTGCTTATGAGTCAAGTCAAGCTTCTTCAAGAGGT 1069
QY 546 LeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGlyIleAlaIleVal 565
DB 1070 CTGTGTGAACCAAGCATGTCAAGTGTCCATGTGTCACCATGTGAGTGTGCCCAA 1129
QY 566 GlySerArgTrpIleAspLeuPheGluProGlyIleValArgArgAlaLeuValIleVal 585
DB 1130 GGTATCAGTTGAAGATTTGTTTCACTGAGTGAAGGCTGCTTGTGTGCTG 1189
QY 586 GlyIleGluIleLeuGluIlePheAlaGlyIleAsnGlyValIleValIleValIleVal 605
DB 1190 GGAATTCAGATCTTCAACAGTTTGTGAATTAACGTTGTTCTGATATACCCACA 1249
QY 606 IleLeuGluIleAlaGlyIleValAlaIleLeuSerIlePheGlyLeuSerSerAlaSer 625
DB 1250 ATTTCTTGAGCAAGCTGAGTGTGCACTTCTTCCAAATTTGCTCAGCTCGGCACA 1309
QY 626 AlaSerIleLeuIleSerSerIleThrIleLeuMetLeuProCysIleGlyPheAla 645
DB 1310 GCATCATCTTATATGATGTTCTCTCACTTACTTAAAGCTTCTTGCATGTTGGCTTGC 1369
QY 646 MetLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyIleProIleLeu 665
DB 1370 ATGCTGCTTATGATCTTCCGGAAGAGGTTTTGCTGCTAGGACAAATCCATCTTG 1429
QY 666 IleAlaSerLeuValIleLeuValIleSerAsnLeuIleAspLeuGlyIleThrAlaHis 685
DB 1430 ATAGCATCTCTATGATCTGCTGTGTGTCATCAATCAATTTGATGGTACACTGACCAT 1489
QY 686 AlaLeuLeuSerThrValSerValIleValIleValIleValIleValIleValIleVal 705
DB 1490 GCTTGTCTCCACCATCAAGTATGCTTACTTCTCTGCTGCTTCTGATAGGATTTGCT 1549
QY 706 ProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgIleValCysIle 725

Db 1550 CCCATCCCAACATTTATGTCAGAGATCTTTCCACCAAGGCTTCGTGCTGTATTT 1609
 Qy 726 AAlleCYsaIaPheThPheTrIleGlyAspIleIleValThrTySerLeuProAl 745
 Db 1610 GCCATTGTGCTTACATTTGATCGGAGATATCATGTCACCTACAGGCTTCCTGTG 1669
 Qy 746 MetLeuSaIaIleGlyLeuAlaGlyValPheSerIleTyAlaValAlaCysLeuIle 765
 Db 1670 ATGCTGAATGCTATTGACCTGCGGGCTTTTCAGCATATATGAGTGTATGCTTGA 1729
 Qy 766 SerPheValPheValPheLeuAlaValProGluThrTyGlyMetProLeuGluValIle 785
 Db 1730 TCCCTTGTGCTTCCTTCTTAAGGTCCTGAGACAAAGGGAGTCCCTTGAAGTATTT 1789
 Qy 786 ThrGluPhePheAlaValAlaGlyAlaGlyLeuAlaIleAlaValAla 800
 Db 1790 ACCGAATTTCTTGACGTTGTGCGAAGCAGCGCTGCAAAAGCC 1834
 RESULT 2
 CC423708/c 911 bp DNA linear GSS 19-MAY-2003
 LOCUS PUH0R49TB.ZM.0.6.1.0_Zea_mays genomic clone ZMMBTA499102,
 DEFINITION genomic survey sequence.
 ACCESSION CC423708
 VERSION CC423708.1 GI:30903798
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 911)
 WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Resnick
 A., Frazer,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.
 Maize Genomics Consortium
 Unpublished
 Other GSSEs: PUH0R49TD
 Contact: Cathy WhiteLaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whiteLaw@tigr.org
 Seq primer: TR
 Class: sheared ends.
 FEATURES
 source
 1..911
 Location/Qualifiers
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBTA499102"
 /clone_lib="ZM.0.6.1.0_KB"
 /note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
 Cot selected genomic DNA library"
 BASE COUNT 242 a 264 c 163 g 242 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3.96e-121 Length: 911
 Score: 1399.00 Matches: 267
 Percent Similarity: 98.90% Conservative: 2
 Best Local Similarity: 98.16% Mismatches: 3
 Query Match: 34.33% Indels: 0
 DB: 29 Gaps: 0
 US-10-051-909-32 (1-800) x CC423708 (1-911)
 Qy 270 SerGlyGluLeuSerLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGlu 289
 Db 818 ACAAGGTGAATGTCCTTCTTCGAAGGTTGAGGTTGAGAGACATTCCTTCATTGAA 759
 Qy 290 GluTyrlleIleGlyProAlaThrGluAlaIleAspLeuValThrAspGlyAspLys 309

Db 758 GAGTACATCATTTGACCTGCGACCGAGCAGCCGATCATCTTTGTACTGACGGATATAG 699
 Qy 310 GluGlnIleThrLeuTyGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLys 329
 Db 698 GACAAATATCACTTTATGCGCCCTGGAAGAGCCAGATCATGATGCTTCGACCTTCCAG 639
 Qy 330 GlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsn 349
 Db 638 GGACCCAGCATGCTTGAAGAGTGTCTTCTTCATCTGTCATGAGGATGATGATGAC 579
 Qy 350 GlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMet 369
 Db 578 CAGAGTGAATCCCTTATGATTCGATTTGTGACACTTTTGTGATGTCATGACATATAG 519
 Qy 370 ProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSer 389
 Db 518 CCTCAAGCTGAGAGATGATGAGACACATTTGTTCCAACTTGGAAAGTATCTTCAGT 459
 Qy 390 ValThrAspGlnHisAlaValAsnGluGlnThrAspGluGluAsnLeuHisArgAspAsp 409
 Db 458 GTCACAGATCAGCATGCGCAAAATGACAGATGGATGAGAAATCTTCATAGGATGAC 399
 Qy 410 GluGluTyrlaSerAspGlyAlaGlyGlyAspTyrgLusAsnLeuHisSerProLeu 429
 Db 398 GAGAGATACGATCTGATGCTGCAAGAGGTGACTATGAGCAATCTTCATAGCCCATTTG 339
 Qy 430 LeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGly 449
 Db 338 CTGTCCAGCGCAGCGCAACAGGTCGGAAGGAGGAGCAATTTGACACATGTCACCGTGA 279
 Qy 450 SerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSer 469
 Db 278 AGTCTTTGACATGAGAGGCAAGGCTCTTAAGGAGGTGAGATGAGTGAAGC 219
 Qy 470 ThrAspIleGlyGlyGlyTyrgGlnLeuAlaTrpLysTrpSerGluTyrgLusAsn 489
 Db 218 ACTGATATCGCTGCGGATGCGACCTGCTTGAATGATGATGAGAGAGAGAT 159
 Qy 490 GlyArgLysGluGlyGlyPheLeuArgValTyrlleuHisGlnGluGlyValProGlySer 509
 Db 158 GGTAGAAAGAGAGGTTGTTCAAAAGAGTCACTTGACCAAGAGGAGGATCTCGGCTCA 99
 Qy 510 ArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPhe 529
 Db 98 AGAAGGGGCTCAATGTTTCACTTCCGCGGTGCGCATGTTCTTGAGGTTAGTGA 39
 Qy 530 ValHisAlaIleAlaLeuValSerGlnSerAlaLeu 541
 Db 38 GTACATGCTGCTGCTTTAGTATGATCAGTCAGACACTT 3
 RESULT 3
 BZ723844 845 bp DNA linear GSS 24-FEB-2003
 LOCUS PUCF160TD.ZM.0.6.1.0_Zea_mays genomic clone ZMMBTA13323,
 DEFINITION genomic survey sequence.
 ACCESSION BZ723844
 VERSION BZ723844.1 GI:28519500
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 845)
 WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Resnick
 A., Frazer,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.
 Maize Genomics Consortium
 Unpublished
 Contact: Cathy WhiteLaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843

Db		618	GAAACAATCAACCTTATTGGGCGCTGAAGAAGGCCAGTCATGATTTGCTCCAGCCTTCCAG	559
Oy		330	GIYPProileMetLeugIYSeRValleuSerLeuAlaSerArgHISglYSeRmetValAsn	349
Db		558	GGACCACGACATGCTTGGAAGTGTGCTTCTCTTGCATCTGCTCAATGGGAGCATGGTAGAC	499
Oy		350	GlInserValProluMeMetAspProIleValThreulePheglYSeRValHIsgluAsnmet	369
Db		498	CAGAGTGTACCCCTTAATGATTCGATGTGCACCTTTTGGTAGTGTCCAGAATAATG	439
Oy		370	ProGlnalaglYglYSeRmetArGSeRThrleuPheProAnPheglYSeRmetPheSer	389
Db		438	CCTCAGCTGAGAGACATATGAGAGCACATTTGTTCCAACTTTGGAAATATGTTCACT	379
Oy		390	ValThraPepGlnHIsalaLYsaSngluIntTpAPeGluLguAsnLeuHISarGAspAsp	409
Db		378	GTCAcAGATCAGCATGCCCCAAAATATGACAGTGGATGATGAAGAATCTTCATAGGATAC	319
Oy		410	GIUGlUTYrAlaSerApBglyVALaglyGLysApTYrGIuaPaPnLeuHISerProleu	429
Db		318	GAGGAGTATCGCATCTGATGTGTCAGAGAGTACATATGAGCAATCTCCATAGCCCATYTG	259
Oy		430	LeuSerArGglnAlathGlyVAlAgLuGLylYsaApIleValHSHisglYHISarGly	449
Db		258	CTGTCCAGGACAGGACACAGTGGCGAAGGAAGACATTTGTGCACCATGTGCACCGGTGA	199
Oy		450	SeralaleuSermetArGArGlnThrleuDeuGluGluGluGlyIYaspGlyValaserSer	469
Db		198	AGTGTGTTGACATATGAGAGCAAAAGCCTCTTAGGGGAGGGGTGAGATGCTGTAGACAGC	139
Oy		470	ThraPPIeglyglYglYTPrlnleuAlatPrlySTPSergLuluYegluGlyglYluAsn	489
Db		138	ACTATATATCGGTGGGGGATGTGACGTTGCTTGGAAATGTGCAGAGAGGAAGTGAAGAT	79
Oy		490	GIYArGlyvsgluGlyglYPhelYArGargValYrleuHISglnGluGlyValProGlySer	509
Db		78	GGTGAAGAAAGGAAGGGTGTCAAAGATCTTACCAACAAAGAGGGATTCCTACTACCTCA	19
Oy		510	ArgArGglYSerlleVal 515	
Db		18	AGAAAGGGCTCAATTGTT 1	
RESULT 5				
B0752737		LOCUS		
DEFINITION		829 bp mRNA linear EST 23-JUL-2002		
LOCUS		MHE41118_E11_I22Z5 wheat salt-stressed root cDNA library Triticum		
ACCESSION		aesttium cdna clone MHE4118_E11_I22, mRNA sequence.		
VERSION		B0752737		
KEYWORDS		B0752737.1 GI:21930519		
SOURCE		EST.		
ORGANISM		Triticum aestivum (bread wheat)		
		Triticum aestivum		
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
		Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;		
		; Triticeae; Triticum.		
REFERENCE		Anderson,O.D., Akhunov,E., Chao,S., Crossman,C., Deal,K., Dvorak,J.,		
AUTHORS		Lazo,G.R., Pham,J., Rauech C.J., Wilson,C. and Woo,J.		
TITLE		The structure and function of the expressed portion of the wheat		
JOURNAL		genome - Salt-stressed root cDNA library		
COMMENT		Unpublished		
		Contact: Olin Anderson		
		US Department of Agriculture, Agriculture Research Service, Pacific		
		West Area, Western Regional Research Center		
		800 Buchanan Street, Albany, CA 94710, USA		
		Tel: 5105595773		
		Fax: 51055955818		
		Email: oanderson@pw.usda.gov		
		Sequences have been trimmed to remove vector sequence and low		
		quality sequence with phred score less than 20		
		Seq primer: SK primer.		
FEATURES		Location/Qualifiers		

TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL Unpublished
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Place: 603011 row: H column: 02.
Location/Qualifiers

FEATURES
source 1..594
/organism="Zea mays"
/mol_type="mRNA"
/cultiivar="B73"
/db_xref="taxon:4577"
/tissue_type="seeding"
/dev_stage="salt stress"
/lab_host="E. coli XL Gold"
/clone_lib="603 - stressed root cDNA library from Wang/Bohnert lab"
/note="Organ: root; Vector: pBluescriptII SK(+); XR; Seeding stressed root cDNA library from Wang/Bohnert lab"

BASE COUNT 210 a 127 c 142 g 115 t
ORIGIN

Alignment Scores:
Pred. No.: 5,13e-82 Length: 594
Score: 981.00 Matches: 197
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 24.07% Indels: 0
DB: 9 Gaps: 0

US-10-051-909-32 (1-800) x A1861088 (1-594)

QY 596 IleaenglyVallleuYrThrPrognilleuGluGlnaGlyValaValalle 615
DB 552 ATAAAGGGTCTTCTGACTATACCCCAAAATTTTGACAGAGCTGGTGGCAGATTAT 533
QY 616 LeuSerlySPheGlyLeuSerSerAlaSerAlaSerlleuLleSerSerleuThr 635
DB 532 CTTTCCAATTTGGTCTAGCTCGGATCAGATCATCTTATCATCTCTCACTACC 473
QY 636 LeuLeuMetLeuProCyslleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg 655
DB 472 TTACTAAATGCTTCTTGCATTGGCTTGGCATGCTGATTGATGATCTTCCGGAAGAGG 413
QY 656 PheLeuLeuLeuGlyThrlleProilleuLleAlaSerleuValilleuValaSer 675
DB 412 TTTTGGCTGAGGACAAATTCATCTGATGATGATCTTCTGATGATCTGATGATCT 353
QY 676 AeuLeuLeuAspLeuGlyThrlleuAlaAlaSerleuSerThrValSerValilleuVal 695
DB 352 AATCAATATGATTTGGGTACTAGCCCAATGCTTCTCTCAACCTCACTGATGATGAT 293
QY 696 TyrPheCysCysPheValMetGlyPheGlyProilleuProAsnilleuCysAlaGlu 715
DB 292 TACTTCTGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 233
QY 716 PheProThrArgValArgGlyLeuCysAlaAlaSerleuPheThrPheTrpilleuGly 735
DB 232 TTTTCCAACGAGGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 173
QY 736 AspilleuValThrlyrSerleuProValMetLeuAsnAlaAlaGlyLeuAlaGlyVal 755
DB 172 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 113
QY 756 PheSerlleuValAlaValaCysleuLleSerPheValPheValPheLeuLysValPro 775
DB 112 TTCAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 53

QY 776 GluThrlySPheGlyMetProLeuGluValilleThrllePhePheAlaValaGly 792
DB 52 GAGACAAAGGGGATGCCCTTGAAGTTATTAACCAATTTCTTGCAGTTGGT 2

RESULT 8
B2819234 817 bp DNA linear GSS 18-MAR-2003
LOCUS PUGD077B ZM 0.61.0 KB Zea mays genomic clone ZMMB7A362B13,
DEFINITION genomic survey sequence.
ACCESSION B2819234
VERSION B2819234.1 GI:29034056
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 817)
White, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick
A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Bennetzen, J.
Maize Genomics Consortium
Unpublished
CONTACT Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..817
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMB7A362B13"
/clone_lib="ZM 0.61.0 KB"
/note="Vector: pCR4-TORO; Site 1: EcoRI; 0.6-1.0 kb high
COR selected genomic DNA library"

FEATURES
source 1..817
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMB7A362B13"
/clone_lib="ZM 0.61.0 KB"
/note="Vector: pCR4-TORO; Site 1: EcoRI; 0.6-1.0 kb high
COR selected genomic DNA library"

BASE COUNT 232 a 145 c 203 g 237 t
ORIGIN

Alignment Scores:
Pred. No.: 1.47e-79 Length: 817
Score: 957.00 Matches: 183
Percent Similarity: 99.46% Conservative: 1
Best Local Similarity: 98.92% Mismatches: 1
Query Match: 23.48% Indels: 0
DB: 29 Gaps: 0

US-10-051-909-32 (1-800) x B2819234 (1-817)

QY 270 SerGlyGluLeuSerleuLeuGlyGlyLeuGluValaGlyGlyAspThrSerlleu 289
DB 262 ACAGTGAATGTCCTTCTTCTGAAAGGTTGAGGTGAGAGACACTTCCATTGAA 321
QY 290 GluTyrillelleGlyProAlaThrGluAlaAspAspLeuValThrAspGlyAspLys 309
DB 322 GAGTACATCATTTGAGCCTGCACAGAGGACCGCATGATCTTGTATCTGACGGTATAG 381
QY 310 GluGlnilleuThrleuTyrGlyProGluGluGlyGlnSerTrpilleuAlaArgProSerly 329
DB 382 GAACAAATCACATTTAAGGAGCTGGAAGAGGACGATGATGATGATGATGATGATGAT 441
QY 330 GlyProilleuMetLeuGlySerValleuSerleuAlaSerArgHisGlySerMetVala 349
DB 442 GAGCCACACATGCTTGAAGAGTGTCTTCTTTCATCTGATGATGATGATGATGATGAT 501
QY 350 GlnSerValProLeuMetAspProilleuValThrleuPheGlySerValHisGluAsnMet 369
DB 502 CAGAGTGAATCCCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 561

QY 370 ProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSer 389
 Db CCTCAAGCTGAGAGAGATGAGAGACATTTCTTCCAACTTGGAAGTATGTTCACT 621
 QY 390 ValThrArgGlnHisAlaIysAsnGluGlnTTPASPGLUGLAsnLeuHisArgAspAsp 409
 Db 622 GTCACGATCAGATCCCAAAATGAGCAGTGGAGAGAGATCTTCATGCGGATGAC 681
 QY 410 GlnGluTyrAlaSerAspGlyAlaGlyIysAspTyrGluAspAsnLeuHisSerProLeu 429
 Db 682 GAGGATGACCATCTGATGCTGAGAGATGACTATGAGACATCTCCATAGCCCATTTG 741
 QY 430 LeuSerArgGlnAlaThrGlyAlaGluGlyIysAspGlnValHisGlnIysArgGly 449
 Db 742 CTGTCCAGCAGAGCAACAGTGTGGAGAGAGAGACATTTGTGACCATGTCACCTGGA 801
 QY 450 SerAlaLeuSerMet 454
 Db 802 AGTGTCTTGACCATG 816
 RESULT 9
 LOCUS BM817360 859 bp mRNA linear EST 05-MAR-2002
 DEFINITION HCl06B05_T3.ab1 HC Hordeum vulgare subsp. vulgare cDNA clone
 [Arabidopsis thaliana], mRNA sequence.
 ACCESSION BM817360 GI:19153374
 VERSION EST
 KEYWORDS
 SOURCE Hordeum vulgare subsp. vulgare
 ORGANISM Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidaeae
 ; Triticeae; Hordeum.
 1 (bases 1 to 859)
 Ozturk,N.Z., Michalowski,C.B., Brazille,S., Borchert,C., Palacio,C.,
 Normand,C., Murphy,C., Kelley,R., Sant,S.A., McLaughlin,H.,
 Frederiksen,M.A. and Bohner,H.J.
 Monitoring large-scale changes in transcript abundance in drought-
 and salt-stressed barley
 Unpublished (2002)
 CONTACT: Mark A. Fredrickson
 Plant Biology
 University of Illinois
 1201 W Gregory Dr., Urbana, IL 61801, USA
 Tel: 2172655473
 Email: bohnertlab@life.uiuc.edu.
 FEATURES
 source
 1..859
 Location/Qualifiers
 /organism="Hordeum vulgare subsp. vulgare"
 /mol_type="mRNA"
 /strain="cv tokak"
 /db_xref="taxon:112509"
 /clone="HCl06B05_T3.ab1"
 /tissue_type="root"
 /dev_stage="3 week old"
 /clone_lib="HC"
 /note="6 and 10 hour drought stress by placing plants on
 moist paper (75% rel. humidity) in light"
 BASE COUNT 179 a 206 c 213 g 260 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 5.53e-77 Length: 859
 Percent: 930.00 Matches: 183
 Percent Similarity: 80.58% Conservative: 41
 Best Local Similarity: 65.83% Mismatches: 51
 Query Match: 22.82% Indels: 3
 DB: 12 Gaps: 1
 US-10-051-909-32 (1-800) x BM817360 (1-859)
 QY 496 PheIysArgValIyLeuHleGlnGluValProGlySerArgArgGlySerIleVal 515

Db |||||.....
 34 TTCCAGAGATTTATTTTCATCAGAGAGTACGCTGGATCAAAACAGGGCTCTATATT 93
 QY 516 SerLeuProGlyGlyValPheGluGlySerGluPheValHisAlaAlaLeu 535
 Db |||||.....
 94 TCTCTT-----GGTGGATGCTGCTGAGAGACAGAGTATATCAAGGCTGCTGACTA 147
 QY 536 ValSerGlnSerAlaLeuPheSerIyGlyLeuAlaGluProArgMetSerAspAla 555
 Db 148 GTTACTCAACAGCCCTTATTTCAGAGAGCTATGATCGGAATTCCGTTGACCACT 207
 QY 556 MetValHisProSerGluValAlaAlaIysGlySerArgTyrIysAspLeuPheGluPro 575
 Db 208 ATGATTCACCCCTCTGAAGCCTCTGCTAAAGGGCTATCTGAGAGAGATCTTTGAGCT 267
 QY 576 GlyValArgAlaLeuLeuValGlyValGlyIleGlnIleGlnIleGlnPheAlaGly 595
 Db 268 GGTTCAGCAGCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 327
 QY 596 IleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaIle 615
 Db 328 ATTAGCGAGATTATGATTTATTTATCTCCCAATATCTTACGCCAGGCTGGGTTGGTATCT 387
 QY 615 eLeuSerIysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThr 635
 Db 388 TCTATCACATPAGGATATGACAGTCACTGCGCTCATTTCTTATGTCACGACGAC 447
 QY 635 rLeuLeuLeuLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg 655
 Db 448 ACTTCTGATGCTTCTCTATAGCTGTTGCTATAGAGCTCATGATATATCAGAGAGAG 507
 QY 655 gPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValIse 675
 Db 508 GAGTTTGTCTGCTAACGACATCCCAATGATGATCTGCTGCTGCTGCTGCTGCTGCTG 567
 QY 675 rAsnLeuIleAspLeuGlyThrIleAlaHisAlaLeuLeuSerThrValSerValIle 695
 Db 568 AAGCTTATTAAGATGAGATCATATTAAGGGGTATCTCCACTGATATGTTGGCT 627
 QY 695 lTyrPheCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGlu 715
 Db 628 TTACTTCTGTTCTTCTGATGAGGCTTGGCCGCTGCTGCTGCTGCTGCTGCTGCTG 687
 QY 715 ePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIle 735
 Db 688 CTTCCNACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 747
 QY 735 yAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGly 755
 Db 748 TGACATATATGCTACCTACACACTCCGCTGATCTCAAAATCTTTCGACCTCGCTGCG 807
 QY 755 lPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeu 772
 Db 808 CTTTGTATCTATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 859
 RESULT 10
 LOCUS BU040051 749 bp mRNA linear EST 26-AUG-2002
 DEFINITION PP_Lea0004L05f Peach developing fruit mesocarp Prunus persica cDNA
 clone PP_Lea0004L05f, mRNA sequence.
 ACCESSION BU040051 GI:22478805
 VERSION EST
 KEYWORDS
 SOURCE Prunus persica (peach)
 ORGANISM Prunus persica
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
 1 (bases 1 to 749)
 Callahan,A., Palmer,M., Main,D., Wing,R. and Abbott,A.
 Peach Model Genome for Rosaceae
 Unpublished
 CONTACT: Abbott, A.
 JOURNAL
 COMMENT

Dept of Genetics and Biochemistry
Clemson University
122 Long Hall, Clemson University, Clemson, SC 29634, USA
Tel: 864 656 3060
Fax: 864 656 6879
Email: aalbert@clemson.edu

Total High Quality bases = 604
Seq primer: TATACGACTCATTAGGAGG
High quality sequence stop: 749.
Location/Qualifiers

FEATURES

source

1. 749
/organism="Prunus persica"
/mol_type="mRNA"
/cultivar="Loring"
/db_xref="taxon:3760"
/clone="PP_LBA004L05F"
/tissue_type="Mesocarp"
/lab_host="E. coli"
/clone_lib="Peach developing fruit mesocarp"
/note="Vector: pBluescript II SK(-); Site 1: EcoRI;
Site 2: XhoI; authority=Prunus persica L. Batsch. The
sequence has been trimmed to remove vector sequence and
contains a minimum of 100 bases of phred value 20 or
above. For more details on library preparation and
sequence analysis go to
http://www.genome.clemson.edu/projects/peach. To order
this clone go to http://www.genome.clemson.edu/orders"

BASE COUNT 167 a 167 c 163 g 252 t
ORIGIN

Alignment Scores:

Pred. No.: 7.02e-77 Length: 749
Score: 928.00 Matches: 172
Percent Similarity: 87.23% Conservative: 33
Best Local Similarity: 73.19% Mismatches: 30
Query Match: 22.77% Indels: 0
Gaps: 13

US-10-051-909-32 (1-800) x BU040051 (1-749)

QY 564 AAlaySgiSerArGTpLySAspleuPhegiuProgiYValArgAlaLeuLeuVal 583
Db 3 TCAGAAAGACCAATCTGGGCTGCTGTTTGAACAGGGGTTAAGCATGATGTTGTT 62
QY 564 GYValAGlylEglnlLeuEnglnlPheAlaGlylLeuEnglyValLeuYrYThr 603
Db 63 GGAATAGGAATCCAGATTCCTCAGACAGTTTCTGCGATTAATAGGGGTTCTGACTACACT 122
QY 604 ProGlnlLeuEnglnlAlaGlylValAlaVallLeuSerlyPheGlyLeuSerSer 623
Db 123 CCTCAATTCCTTGAAGATGACAGGTGTGAAGTCTTCTGAAGACTTGGGTCCTGATCA 182
QY 624 AAlaSerAlaSerlLeuLeuSerSerLeuThrLeuLeuMetLeuProCyAlleGly 643
Db 183 GAGTCTTATCCTTCTCATATAGTCAATCAACACTTTGTTATCTCTCTGTTATGACC 242
QY 644 PheAlaMetLeuLeuMetAspleuSerGlyAGTAgPheLeuLeuEnglyThrlePro 663
Db 243 ATAGCCATAAAGCTATGATGTCTCTGTGTAAAGAGACGCTGCTACTGCTACAAATTCCT 302
QY 664 lLeuLeuAlaSerLeuVallLeuValSerAlaSerAlaSerAlaSerLeu 683
Db 303 GGTGTGTAGTACACTATCTATCTATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 362
QY 684 AAlaAlaAlaLeuLeuSerThValSerValleVallyrPheCyCyAsPheValMetGly 703
Db 363 CTTTATGGGGCTTATACAGTCAATTTGTGTATGCTATTTCTGCTGCTGCTGCTGCTGCTG 422
QY 704 PheGlyProleleProAsnleleuCyAlaGlylLeuPheProThrArgValArglyLeu 723
Db 423 TATGGGCGCAATTCCAATATCTCTGCTGAGATCTTTCACAAAGGGGTAGTGGCTCTC 482
QY 724 CyslleAlaAlaCysAlaAlaPheThrPheThrPheleGlyAspAlaAlaValThrlySerleu 743

Db 483 TGCATTCGCATCTCGCCCTAGTGTACTGATTCGGGACATTATGCTACACTACACTA 542
QY 744 ProValMetLeuAlaAlaGlylLeuAlaGlylValPheSerllyrYAlaValAlaCys 763
Db 543 CCTGTGCTGCTTGAATGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 602
QY 764 LeuLeuSerPheValPheValPheLeuLeuValProGlnlThrlySgLyMetProLeuGlu 783
Db 603 GTCAATTCCTTATATATATCTCTGTAAGTTCAGAAACCAAGGATGCCCCCTTGAA 662
QY 784 ValleThrGluPhePheAlaValAlaGlylAlaGlylAlaAlaAla 798
Db 663 GTCAATTCATGATTCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 707

RESULT 11

CA502649

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. 688

/organism="Triticum aestivum"

/mol_type="mRNA"

/cultivar="Chinese Spring"

/db_xref="taxon:4565"

/clone="WHE4338_A09_A18"

/tissue_type="Anther"

/dev_stage="Meiotic stages pre-meiosis-metaphase I"

/lab_host="E. coli DH10B"

/clone_lib="Wheat meiotic anther cDNA library"

/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; Plants

were grown in a glasshouse. Anther meiotic stage was

determined by removing anthers from individual primary

florets. One anther was sacrificed for microscopic staging

and if determined to be between (and including) meiotic

stages pre-meiosis and metaphase I, the remaining two

anthers were collected and pooled for library

construction. The tissue, total RNA, and poly(A) RNA were

prepared, cDNA synthesised, and directionally ligated into

pSPORT1 by Tim Sutton in the P Langridge Lab at the

Department of Plant Science, University of Adelaide, Waite

Campus, Australia. Average insert size 1.5kb. Plasmid DNA

preparations and DNA sequencing were performed in the OD

Anderson Lab (all other authors)."

BASE COUNT

ORIGIN

Alignment Scores:

138 a 167 c 165 g 218 t

DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES	SOURCE
Db														
380	ATGTTCTTCCTTACTGATGATGTTTTCGGAGTGTGTTACGACGAGATCCAACTCGGAG													
Qy														
220	11ethetleuglyValleu1a11eProSerleuphePhePheglyLeuThrIlePheTyr													
Db														
440	TGTGATGCTCGGCGTTCCTTCAATTCACCTTCCTTCAATTCATTTTCACGATTTTTC													
Qy														
240	LeuPProGluSerProArgTrpLeuValSerIlysgIlyArgMetAlaGluAlaIlyLeuVal													
Db														
500	TTGCCGAGTCTCCGAGATGGCTGTTCGACAAAGAAATGTTGAAGCAGACAGGTT													
Qy														
260	LeuGlnIlyLeuArgIlyLysAspAspValSerGlyIguLeuSerLeuLeuLeuGlnIly													
Db														
560	TTGCAGAGATTACGTCGTAGAGAAATTTTCAGAGTGAAGATGGCTTACTGTTGAAGA													
Qy														
280	LeuGlnValIlyGlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAla													
Db														
620	CTTGAAGTTGGGGGAGACACACTTTTGAAGATCACATATATAGGCCCT-----GCA													
Qy														
300	AlaAspAspLeuValThrAspGlyAspIlyS 309													
Db														
671	CATATCAACCCATCACAAGCTGATATAA 700													
RESULT 13														
CA502490		703 bp mRNA linear EST 14-NOV-2002												
LOCUS														
DEFINITION	WHE4048_B04_D08ZT Wheat meiotic anther cDNA library Triticum aestivum cDNA clone WHE4048_B04_D08, mRNA sequence.													
ACCESSION	CA502490													
VERSION	CA502490.1													
KEYWORDS	EST.													
SOURCE	Triticum aestivum (bread wheat)													
ORGANISM	Triticum aestivum													
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae													
AUTHORS	1 (bases 1 to 703)													
TITLE	Anderson, O.D., Chao, S., Crossman, C., Langridge, P., Lazo, G.R., Pham, J., Rauech, C.J., Sutton, T., Woo, J., and Wilson, C.													
JOURNAL	The structure and function of the expressed portion of the wheat genomes - Meiotic anther cDNA library													
COMMENT	Unpublished													
CONTACT	Contact: Olin Anderson													
US DEPARTMENT OF AGRICULTURE, AGRICULTURE RESEARCH SERVICE, PACIFIC														
WEST AREA, WESTERN REGIONAL RESEARCH CENTER														
800 BUCHANAN STREET, ALBANY, CA 94710, USA														
TEL: 5105595773														
FAX: 5105595818														
EMAIL: oanderson@pw.usda.gov														
SEQUENCES HAVE BEEN TRIMMED TO REMOVE VECTOR SEQUENCE AND LOW														
QUALITY SEQUENCE WITH PHRED SCORE LESS THAN 20														
SEQ PRIMER: T7 PRIMER.														
LOCATION/QUALIFIERS														
1..703														

Campus, Australia. Average insert size 1.5Kb. Plasmid DNA
 preparations and DNA sequencing were performed in the OD
 Anderson lab (all other authors). "

BASE COUNT	139 a	183 c	165 g	216 t
ORIGIN				

Pred. No.:	2.63e-76	Length:	703
Score:	921.50	Matches:	189
Percent Similarity:	86.70%	Conservative:	13
Best Local Similarity:	81.12%	Mismatches:	24
Query Match:	22.61%	Indels:	7
DB:	14	Gaps:	2

US-10-051-909-32 (1-800) x CA502490 (1-703)

Qy	526	GlySerGluPhe-----ValHsaIaaIaaIaaLeuVaISerGlnSer	533
Dd	3	GGTCGGAAATTTCCGGGTGCACCGAGCGCTCCGACGCGTGTGGTAAGCACACTCG	62
Qy	540	AlaLeuPheSerLySGlyLeuValGluProArgMet---SerAspAlaAlaMetValHis	556
Dd	63	GCCTTTACTCCAAGGATTATTGAAGAAGGTATAGGGGGTGTCGCCCATGATTCAT	122
Qy	559	ProSerGluValAlaAlaLysGlySerArgTrpIlyAspLeuPheGluProGlyValArg	578
Dd	123	CCATTGGAGGAGCAGCTCCCMAAGTTCAATCTGGAAAGATCTGTTGAACTGGTGTAGG	182
Qy	579	ArgAlaLeuLeuValGlyValGlyTyrLegIlnleuGlnGlnIlePheAlaGlyTlaAsnGly	599
Dd	183	CGTCATTGTTGTCGGCGCTTGGAAATTCAGATGCTTCAGCAGTTTCTGGATTAATAGA	24
Qy	599	ValLeuTyrrTrThrProGlnIleLeuGlnAlaGlyValAlaValIleLeuSerIys	618
Dd	243	GTTCCTACTAATACTCTCAAAATTCGTGAGCAGCTGTGGCGGTCTCTTCCAAAT	302
Qy	619	PheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThleuLeuMet	633
Dd	303	CTTGGGCTCAGTTCAAGATCAGATCAGATTCATCTTGACATCTTCACACCCTTACTATG	362
Qy	639	IeuProCysIIlgIYPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeu	658
Dd	363	CTCCACAGCATTTGGTGTACCATGAGACTTATGATATATCTGAAGAAGGTTTCTGTA	422
Qy	659	LeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValaSerAsnLeuIle	678
Dd	423	CTTAGGACAATTCCTCCATCTTGATACATCCCAATTAATTTGGTTGGTGTCCAAATGATC	482
Qy	679	AspLeuGlyThrIleuAlaHisAlaLeuLeuSerThrValaSerAlaIleValTyrPhcCys	699
Dd	483	AACCTTAGAGGAGTGGTCCACGCTGGCTCTCCACAGTTAGGTCATTTGCTACTCTGC	542
Qy	699	CysPheValaMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThr	718
Dd	543	TGCTTTGTGATGGGCTTTGGCCCGATCCCAACATTTATGAGCAGAGATTTTCCCCACC	602
Qy	719	ArgValaArgGlyLeuCysAlaIleAlaIleCysAlaIlePheTrpIleGlyAspIleIle	733
Dd	603	AGAGCCCGGGTGTCTGCATCGCTATTTGGCGCTCACAAATTCGTGATTTGGACATTAAT	662
Qy	739	ValThrTySerLeuProValaMetLeuAsnAlaIleGly	751
Dd	663	GGTACTACAGCTGCTGTGATGCTGAATGCTATTGCT	701

RESULT 14
 CB981615 847 bp mRNA linear EST 01-MAY-2007
 LOCUS CAB70005 .iiiaf_A05 Cabernet Sauvignon Berry Post-Version - CAB7
 DEFINITION vitis vinifera cDNA clone CAB70005_iiiaf_A05 5', mRNA sequence.
 ACCESSION CB981615
 VERSION CB981615.1 GI:30304821
 KEYWORDS EST.
 SOURCE Vitis vinifera

ORGANISM *Vitis vinifera*
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; Vitaceae; Vitis
 1 (bases 1 to 847)
 Authors Goes da Silva, F., Iandolo, A., Lim, H., Baek, J., Jones, K. and Cook
 D.
 TITLE Expressed sequence tags from cabernet sauvignon berries at various
 developmental stages
 JOURNAL Unpublished
 COMMENT Contact: Douglas Cook, PhD
 CABIS Genome Facility
 UC Davis, Plant Pathology
 One Shields Ave, Davis, CA 95616, USA
 Tel: 530 754 6561
 Fax: 530 754 6617
 Email: drcoc@ucdavis.edu
 Seq primer: ACGTACCGACATATGCC.
 Location/Qualifiers
 1. 847
 /organism="Vitis vinifera"
 /mol_type="mRNA"
 /cultivar="Cabernet Sauvignon"
 /db_xref="taxon:29760"
 /clone="CAB70005.111af_A05"
 /sex="Hermaphrodite"
 /dev_stage="Post-Veraison, 13-19 drix"
 /lab_host="PH5alpha"
 /clone_lib="Cabernet Sauvignon Berry Post-Veraison - CAB7"
 /note="Organ: Berry; Vector: pDNR; Site: 1; Site 2:
 5'11; CAB7 is a cDNA library of Cabernet Sauvignon Clone 8
 berries. Samples were collected post-veraison from
 field-grown vines during stage III of berry growth at 89
 days after full bloom. Berries soluble solid concentration
 ranged between 18-19 brix. Sampled vines were located at
 the University of California, Davis, Experimental
 Vineyard. cDNAs were made by oligo-dT priming and direct
 cloning. 5' and 3' adaptors were used in cloning as
 follows: 5'-AAGCAGTGTATCATCAGACGAGGCGCATTCAGCGCGG-3' and
 5'-ATTCTAGAGCCGAGCGCGCCACATG-dt(30)NN-3'. Library was
 constructed using the Clontech Creator SMART kit and
 size-selected to contain the 0.5-3 kb size fraction."

BASE COUNT 174 a 192 c 185 g 296 t

ORIGIN

Alignment Scores:
 Pred. No.: 1.13e-75 Length: 847
 Score: 916.00 Matches: 173
 Percent Similarity: 88.99% Conservative: 29
 Best Local Similarity: 76.21% Mismatches: 25
 Query Match: 22.48% Indels: 0
 DB: 14 Gaps: 0

US-10-051-909-32 (1-800) x CB981615 (1-847)

QY 573 Phegiuproglyvvalararalaleuvalaiglyvalaiglyleuglnleugln 592
 Db 3 TTTAAACGAGAGCTCAAGATGATGCTGTTGGGAGGAGGATTCAGATCTTCAAGCAG 62
 QY 553 Phealaglyleanglyvalaiglyleuglyleuglyleuglyleuglyleugly 612
 Db 63 TTCTCTGGCATTAATGGGGTCTGTGATCTACCTCCCAATCTTGAAGAGCAGAGTGT 122
 QY 613 AlaValIleleuSerIlePheGlyleuSerSerIleSerIleleuIleSerSer 632
 Db 123 GGAAGTATCTTTCAACATAGCATATGATCTTCACTTCTCTCTTCTTCTAGTCA 182
 QY 633 LeuThrThrleuLeuMetleuProCysIleGlyPheAlaMetleuLeuMetAspLeuSer 652
 Db 183 ATGACAACTTGTATGCTTCTTGTGATGATGCTGTTCCTTGAAGCTCAGATATTTCT 242
 QY 653 G1ATGATGPhelLeuLeuGlyThrIleProIleLeuIleAlaSerleuValIleLeu 672

Db 243 GGTAGAAAGAGTGTCTGCTCAGCAATCCCTGCTCATATGATGACCTTCATTTCTT 302
 QY 673 ValValSeranleuIleAspLeuGlyThrleuAlaHisAlaLeuSerThrValSer 692
 Db 303 GTCTTGGAAGTTTATGTAATATGGTGATGTGTGATGTGCAATCTCAGCTGATG 362
 QY 693 ValIleValIlePheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCys 712
 Db 363 GTCAATCATCTACTTCTGTGTGCTTGTGATGGATTTGGCCCAAGCCCAATATCTGCT 422
 QY 713 AlaGlyIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPhe 732
 Db 423 GCGAGATCTTTCCACCCGCTGCTCAGGCTTATGATTCATTTGTCATCTCTCA 482
 QY 733 TrpIleGlyAspIleIleValThrTyrSerleuProValMetLeuAsnAlaIleGlyLeu 752
 Db 483 TGGATGTGATCATCTATGATTAACCTTACATGCTGCAATGATGATCCTCTGTGGCCTT 542
 QY 753 AlaGlyValPheSerIleTyrAlaValAlaCysleuIleSerPheAlaPheValPheLeu 772
 Db 543 GCTGTGTCTTTGGCATGTACGCTGTGTGTGCTCATCTCATGGGTGCTTCTTCTG 602
 QY 773 LysValProGlyThrLysGlyMetProleuGlyValIleThrGlyPhePheAlaValGly 792
 Db 603 AAGTCCCAAGAAACCAAGGATGCTTGAAGTTATCTGTGAGTTCTTGTGCTGTGT 662
 QY 793 AlaLysGlnAlaAlaAlaLys 799
 Db 663 GCAAGTGTGTGAGCAAGAGAG 683

RESULT 15
 BM325261 572 bp mRNA linear EST 04-JAN-2002
 LOCUS P1C1_42_E11.b1.A002 pathogen-infected compatible 1 (P1C1) Sorghum
 DEFINITION
 ACCESSION BM325261
 VERSION BM325261.1 GI:18064398
 KEYWORDS EST.
 SOURCE Sorghum bicolor (sorghum)
 ORGANISM Sorghum bicolor
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Sorghum.
 1 (bases 1 to 572)
 Authors Cordonnier-Pratt, M.-M., Gingle, A., Fang, G.C., Dean, R., Wing, R.,
 Sudman, M. and Pratt, L.H.
 An EST database from Sorghum: plants infected with a compatible
 pathogen
 JOURNAL Unpublished
 COMMENT Contact: Cordonnier-Pratt MM
 Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 583 0210
 Email: mmpratt@uga.edu
 Sequences have been trimmed to exclude PolyA, vector, and regions
 below phred quality 16. The threshold for highest quality sequence
 is 20. Three-prime sequences, which are obtained with PolyMix or
 T7 sequencing primer, are presented as the reverse complement.
 Seq primer: JEN REV
 High quality sequence stop: 543
 POLYA=No.

FEATURES
 source
 1. 572
 /organism="Sorghum bicolor"
 /mol_type="mRNA"
 /cultivar="BTx623"
 /db_xref="taxon:4558"
 /tissue_type="Leaves"
 /dev_stage="4-week-old seedlings infected with
 Colletotrichum graminicola"
 /clone_lib="Pathogen-infected compatible 1 (P1C1)"

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 5, 2004, 21:45:08 ; Search time 648 Seconds
(without alignments)
4273.180 Million cell updates/sec

Title: US-10-051-909-32

Perfect score: 4075

Sequence: 1 IRSGSWLAVQTFPFDLDR.....PLEVTEFFAVGAKQAQAAKA 800

Scoring table:

BLOSUM62	
Xgapop 10.0, Xgapext 0.5	
Ygapop 10.0, Ygapext 0.5	
Fgapop 6.0, Fgapext 7.0	
Delop 6.0, Delext 7.0	

Searched: 2263443 seqs, 1730637950 residues

Total number of hits satisfying chosen parameters: 4526886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

MODEL=frame.p2n.model -DEV-xih
-O=/cgn2_1/USPTO-SPool/US10051909/runat_05012004_094745_22401/app_query.fasta_1.967
-DB=Published Applications_NA -GFM=faetap -SUFFIX=rmpb -MINMATCH=0.1
-LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pcp -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pcp -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10051909@cgn.1.1.418@runat_05012004_094745_22401
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELXT=7

Database :

Published Applications_NA.*

1:	/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2:	/cgn2_6/ptodata/2/pubpna/PTCT_NEW_PUB.seq.*
3:	/cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4:	/cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5:	/cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6:	/cgn2_6/ptodata/2/pubpna/PTCTUS_PUBCOMB.seq.*
7:	/cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9:	/cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10:	/cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11:	/cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
14:	/cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
15:	/cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
16:	/cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
17:	/cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
18:	/cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
------------	-------	--------------------	----	-------------

1	4075	100.0	2777	14	US-10-051-909-31	Sequence 31, Appl
2	4059	99.6	2908	12	US-10-310-154-350	Sequence 350, App
3	3517	86.3	2824	14	US-10-051-902-1	Sequence 1, Appl
4	3517	86.3	2824	14	US-10-051-909-1	Sequence 1, Appl
5	2674	65.6	2601	14	US-10-051-902-7	Sequence 7, Appl
6	2674	65.6	2601	14	US-10-051-909-7	Sequence 7, Appl
7	2584	63.4	2190	10	US-09-938-842A-1315	Sequence 1315, Ap
8	2275	55.8	2205	10	US-09-938-842A-2254	Sequence 2254, Ap
9	1670.5	41.0	1692	14	US-10-051-902-9	Sequence 9, Appl
10	1670.5	41.0	1692	14	US-10-051-909-9	Sequence 9, Appl
11	1424	34.9	1487	14	US-10-051-902-13	Sequence 13, Appl
12	1424	34.9	1487	14	US-10-051-909-13	Sequence 13, Appl
13	943	23.1	1009	14	US-10-051-902-15	Sequence 15, Appl
14	943	23.1	1009	14	US-10-051-909-15	Sequence 15, Appl
15	587.5	14.4	1374	12	US-10-369-493-47011	Sequence 47011, A
16	562	13.8	870	14	US-10-051-902-5	Sequence 5, Appl
17	562	13.8	870	14	US-10-051-909-5	Sequence 5, Appl
18	539	13.2	1395	10	US-09-974-300-907	Sequence 907, App
19	534	13.1	2127	12	US-10-310-154-354	Sequence 354, App
20	533	13.1	1386	12	US-10-369-493-47058	Sequence 47058, A
21	505	12.4	1422	12	US-10-369-493-46784	Sequence 46784, A
22	496	12.2	1853	14	US-10-051-902-23	Sequence 23, Appl
23	496	12.2	1853	14	US-10-051-909-23	Sequence 23, Appl
24	494.5	12.1	2017	14	US-10-051-902-21	Sequence 21, Appl
25	494.5	12.1	2017	14	US-10-051-909-21	Sequence 21, Appl
26	493	12.1	1482	10	US-09-938-842A-795	Sequence 795, App
27	491.5	12.1	1914	14	US-10-051-902-19	Sequence 19, Appl
28	491.5	12.1	1914	14	US-10-051-909-19	Sequence 19, Appl
29	489	12.0	1566	10	US-09-938-842A-1769	Sequence 1769, Ap
30	489	12.0	1872	14	US-10-051-902-27	Sequence 27, Appl
31	489	12.0	1872	14	US-10-051-909-27	Sequence 27, Appl
32	486.5	11.9	1449	12	US-10-369-493-41339	Sequence 41339, A
33	475	11.7	1644	10	US-09-938-842A-3120	Sequence 3120, App
34	473.5	11.6	1395	12	US-10-369-493-46891	Sequence 46891, A
35	469.5	11.5	2051	12	US-10-310-154-356	Sequence 356, App
36	465	11.4	1260	12	US-10-369-493-46835	Sequence 46835, A
37	465	11.4	1931	12	US-10-310-154-355	Sequence 355, App
38	465	11.4	2063	14	US-10-051-902-33	Sequence 33, Appl
39	463.5	11.4	2089	14	US-10-051-902-25	Sequence 25, Appl
40	463.5	11.4	2089	14	US-10-051-909-25	Sequence 25, Appl
41	462	11.3	1839	12	US-10-369-493-25706	Sequence 25706, A
42	459	11.3	285	9	US-09-923-876-6145	Sequence 6145, Ap
43	459	11.3	285	12	US-09-923-876-6145	Sequence 6145, Ap
44	453.5	11.1	1905	9	US-09-734-569-171	Sequence 171, App
45	451	11.1	1377	12	US-10-369-493-39348	Sequence 39348, A

ALIGNMENTS

RESULT 1
US-10-051-909-31
Sequence 31, Application US/10051909
Publication No. US20020199217A1
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Helen, Jarvis, Tim
APPLICANT: Hitz, Bill
APPLICANT: Kinney, Tony
APPLICANT: Timney, Scott
TITLE OF INVENTION: Plant Sugar Transport Proteins
FILE REFERENCE: B81163 US CIP
CURRENT APPLICATION NUMBER: US/10/051, 909
CURRENT FILING DATE: 2002-01-17
PRIOR APPLICATION NUMBER: 60/083, 044
PRIOR FILING DATE: April 24, 1998
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Microsoft Office 97
SEQ ID NO 31
LENGTH: 2777
TYPE: DNA
ORGANISM: Zea mays
US-10-051-909-31


```

Db      826 ATGCTTGAGTCTGCGATACCTTCTTCTTTGTTGACATATTTATCT 885
Qy      241 ProGluSerProArgTrpLeuValSerIleGlyAlaGlyMetAlaGluAlaIleValSerIle 260
Db      886 CCGATATCTCCCAAGATGCTCGTTAGCAAAAGCTCGATGCGAGAGCAAAAAGGTGTTG 945
Qy      261 GluIleValSerArgGlyLeuAspAspValSerIleGlyIleLeuSerIleLeuLeuLeu 280
Db      946 CAAAGCTTACGGGGGAGAAAGACATGTTCTAGGTGTAATGTCCTTCTTCTGCAAGGGTTG 1005
Qy      281 GluValIleGlyIleAspThrSerIleGlyIleIleGlyProAlaThrGluAla 300
Db      1006 GAGGTGGAGAGACACTTCATTTAGAGAGTACATCATTTGACCTGCGACCGAGGAGCC 1065
Qy      301 AspaPleuValThrAspGlyAspIleGlyIleIleThrLeuTrpGlyProGluLeu 320
Db      1066 GATGATCTTGTATACGACGGTATAGGAACAATCACACTTTATGGGCTGAGAAAGGC 1125
Qy      321 GluSerTrpIleAlaArgProSerIleGlyProIleMetLeuGlySerValLeuSerLeu 340
Db      1126 CAGTATGATGATTCCTGACCTTCCAAAGGAGCCAGATGCTTGGAAAGTGTCTTCTTCTT 1185
Qy      341 AlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThr 360
Db      1186 GCATCTCGTCATGGAGCATGTGTGAACCAAGATGTAACCTTATGATCCGATTTGACACA 1245
Qy      361 LeuPheGlySerValHisGluAsnMetProGlnAlaGlyIleSerMetArgSerThrLeu 380
Db      1246 CTTTGTGTAGTGTCCATGAGATATGCTCTCAAGCTGGAGGAATGATGGAGCAATGG 1305
Qy      381 PheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaIleAsnGlnIleTrp 400
Db      1306 TTTCCAACTTTGGAAAGTATGTTCAAGTGCACAGATCCAGATCCCAAAATATAGCAATGG 1365
Qy      401 AspGluGluAsnLeuHisArgAspAspGluGluIleValSerAspGlyAlaGlyIleAsp 420
Db      1366 GATGAGAGAAATCTTCATGAGGATGACGAGAGTACGATGTCATGATGTCAGAGAGTGC 1425
Qy      421 TyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluIleVal 440
Db      1426 TATGAGGACAAATCTCCATAGCCCATTTGCTGTCCAGGCAAGGTCGCGAAGGGAAG 1485
Qy      441 AspiLeuValHisHisGlyHisArgIleSerAlaLeuSerMetArgArgIleThrLeuLeu 460
Db      1486 GACATTTGTGACACATGATGACCGTGGAATGCTTTAGCATAGAAAGGAAAGCCCTTTA 1545
Qy      461 GlyIleGluGlyIleAspGlyValSerSerThrAspIleGlyIleGlyIleTrpGlnLeuAlaTrp 480
Db      1546 GGGGAGGGGTGAGATGCTGTGAGCAACCTGATATCGGTGGGGGATGGCAGCTTCTGG 1605
Qy      481 LysTrpSerGluLeuGluGluIleGluAsnGlyAlaGlyIleGlyIlePheIleValArgValIle 500
Db      1606 AAATGTCAGAGAAAGAGTGTAGATGTAGAAAGAGAGGTTCCTCAAAAGAGTCTAC 1665
Qy      501 LeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyIle 520
Db      1666 TTGCACCAAGAGAGGAGTCTGCTGCTCAGAAAGGGGCTCAATGTTTCACTTCCCGTGT 1725
Qy      521 GlyAspValPheGluGlySerGluPheValHisAlaAlaIleValSerGlnSerAla 540
Db      1726 GCGCATGTTCTTGAAGGTAGTGAATTTGTATCATGCTGCTTGTAGTACAGCAGCA 1785
Qy      541 LeuPheSerIleGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSer 560
Db      1786 CTTTCTCAAAAGGGTCTGCTGACCAACCGCATGTCAAGTGTCCATGATGTTCAACCATCT 1845
Qy      561 GluValAlaAlaIleValSerArgTrpIleAspLeuPheGluProGlyValAlaArgAla 580
Db      1846 GAGGTAGCTGCCAAAGGTTCACTGTGGAAGATTTCTTGAACCTGTGAGTGGGGGTGCC 1905
Qy      581 LeuLeuValGlyValGlyIleGlnIleLeuGlnIlePheAlaGlyIleAsnGlyValLeu 600

```

```

Db      1906 CTGTATAGTGGGTGTTGGAATTCAGATCCTTCAACAGTTTGTGGAAATTAACGGTGTCTG 1965
Qy      601 TyrTrpThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerIlePheGly 620
Db      1966 TACTATACCCCAAAATTTCTTGACCAAGCTGTGTGGAGTTATTTCTTCCAAATTTGGT 2025
Qy      621 LeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrTrpLeuMetLeuPro 640
Db      2026 CTCAGCTGGAGTACGATCCATCTTATGATCGTTCTCCTACATCTTACTTACTTACTTCT 2085
Qy      641 CysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGly 660
Db      2086 TGCAATGGCTTTGGCATGCTCTTATGATCTTCCGGAAGAGTTTCTGCTAGGC 2145
Qy      661 ThrIleProIleLeuIleAlaSerLeuValIleLeuValIleSerAspLeuIleAspLeu 680
Db      2146 ACAATTCAAATCTTGATAGCATCTGATGTAATCCGTGTGTGTCCAATCTAATGATTTG 2205
Qy      681 GlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValIleValIleValIleVal 700
Db      2206 GGTACACTAGCCCATGCTTCTCTCCAGCTCAGTGTATCTGCTACTTCTGCTGCTTCT 2265
Qy      701 ValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgVal 720
Db      2266 GTTATGGATTTGGTCCCATCCCAACATTTATGTGAGAGATCTTCCAAACAGGGGT 2325
Qy      721 ArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThr 740
Db      2326 CGTGCCCTGTGATTTGCCATTTGCTTTCATTTGATGATGATGATGATGATGATGATGATG 2385
Qy      741 TyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAla 760
Db      2386 TACAGCTTCTCTGATGCTGAATGCTATTTGATGATGATGATGATGATGATGATGATGATG 2445
Qy      761 ValValCysLeuIleSerPheValPheValPheLeuIleValIleProGluThrIleGlyMet 780
Db      2446 GTCGATCTGATTTCTTGTGTGTGCTTCTTAAAGTCCCTGAGACAAAGGGGATG 2505
Qy      781 ProLeuGluValIleThrGluPhePheAlaValAlaGlyAlaIleValIleValIleValIle 800
Db      2506 CCCCTGAGGTATTAACGAATCTTGTGACATGATGATGATGATGATGATGATGATGATG 2565

RESULT 3
US-10-051-902-1
; Sequence 1, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hite, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 2824
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURES:
; NAME/KEY: unsure
; LOCATION: (29)
; NAME/KEY: unsure
; LOCATION: (622)
; NAME/KEY: unsure
; LOCATION: (636)
; NAME/KEY: unsure
; LOCATION: (638)
; NAME/KEY: unsure

```



```

Db 2038 AGAAGTTTGGCTGCTAGGACCAATTCATCTGATGATCTGATCTGCTGTT 2097
Qy 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaIleLeuSerThrValSerVal 693
Db 2098 GTGTCAATCTAATGATTTGGTACACTAGCCCAAGCTTGTCTCCACCACTAGTGT 2157
Qy 694 IleValIlePheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
Db 2158 ATGCTCTACTTGTGCTGCTCTTATGGATTTGGTCCCATCCCAACTTTTATGTGCA 2217
Qy 714 GluIlePheProThrArgValArgGlyLeuGlyIleAlaIleCysAlaPheThrPheTrp 733
Db 2218 GAGATCTTCCACACAGGGTGTGGCTCTGTATTCATTCGATTTGACCTTATTCATTCGG 2277
Qy 734 IleGlyAspIleIleValIleThrIleSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
Db 2278 ATGGAGATATCATGTATGCACTACAGCCCTTCTGTATGCTGAATGCTAATGGACTGACG 2337
Qy 754 GlyValPheSerIleIleValIleValIleCysLeuIleSerPheValPheValPheLeuLys 773
Db 2338 GGTGTTTTCAGCATATATGCACTGATGCTTGTGATTTCTTGTGCTTCTTCTTATG 2397
Qy 774 ValProGluThrIleGlyMetProLeuGluValIleThrGluPhePheAlaValGlyVal 793
Db 2398 GTCCCTGAGACAAAGGGATGCCCCCTTGAGGTTATTAACGAATCTTTGCAAGTTGATGCG 2457
Qy 794 LysGluAlaIleAlaIleValAla 800
Db 2458 AAGCAAGCGGCTGCAAAAGCC 2478

```

RESULT 4

```

US-10-051-909-1
; Sequence 1, Application US/10051909
; Publication No. US2002019217A1

```

GENERAL INFORMATION:

```

; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tinney, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO. 1
; LENGTH: 2824

```

TYPE: DNA

```

; ORGANISM: Zea mays
; FEATURE:

```

```

; NAME/KEY: unsure
; LOCATION: (129)
; NAME/KEY: unsure
; LOCATION: (622)
; NAME/KEY: unsure
; LOCATION: (636)
; NAME/KEY: unsure
; LOCATION: (638)
; NAME/KEY: unsure
; LOCATION: (669)
; NAME/KEY: unsure
; LOCATION: (771)
; NAME/KEY: unsure
; LOCATION: (822)
; NAME/KEY: unsure
; LOCATION: (856)
; NAME/KEY: unsure
; LOCATION: (889)
; NAME/KEY: unsure

```

```

; LOCATION: (896)
; NAME/KEY: unsure
; LOCATION: (944)
US-10-051-909-1

```

Alignment Scores:

```

Pred. No.: 0 Length: 2824
Score: 3517.00 Matches: 684
Percent Similarity: 95.18% Conservative: 27
Best Local Similarity: 91.57% Mismatches: 34
Query Match: 86.31% Indels: 2
DB: 14 Gaps: 2

```

```

US-10-051-909-32 (1-800) x US-10-051-909-1 (1-2824)

```

```

Qy 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlyTrp 75
Db 238 ATGGGGGGCGCGAGATGCTGCAATCGCGCCCTCTATCGGCAACTGTGCGAGGCTGG 297
Qy 76 AspaAsnAlaThrIleAlaIleAlaValLeuTrpIleLysLysGluPheGlnLeuGlnAsn 95
Db 298 GACAAATGCAAAATTCCTGAGCCGCTCTGTACTTAAGAAAGAAATTCACCTGACAGC 357
Qy 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db 358 GAGCCTGTATGCAAGGCTCATCTCCTGATGCTGCTCATTTGAGGGCAAGCATCATCA 417
Qy 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
Db 418 ACATCTCCGGGGCCCAAGGCTGACTGCTGTGTGAGGGCCCAATGCTGTGCTCGCT 477
Qy 136 IleLeuTrpPhePheSerGlyLeuIleMetLeuTrpSerProAsnValIleValLeuLeu 155
Db 478 GTCTTACTTCTGCAATGAGGCTGTGATGCTTTGGGCGCAATGTGTACTTGTCTC 537
Qy 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaIleThrLeuValProLeuTrp 175
Db 538 CTCGCAAGGCTCATGTATGAGGCTGATCGGTTTGGGGGCACACACTTGTCTCTAC 597
Qy 176 IleSerGluIleAlaProSerGluIleArg--GlyLeuLeuAsnThrLeuProGlnPhe 194
Db 598 ATCTCCGAACCTGACCGCACAGAAATCTTGGGCGTGMTNGAACAGCTTGCACAGTTC 657
Qy 195 SerGly---SerGlyGlyMetPheLeuSerTrpCysMetValPheGlyMetSerLeuSer 213
Db 658 ATTGGGCTCAAGGAGGATTTCTCTCTACTGCAATGAGTTTGGGATTCCTCATG 717
Qy 214 ProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhe 233
Db 718 CCCAACTGATTTGAGGCTCATGCTTGAAGTTCTGTCGATCCCGTCACTTATATCTTT 777
Qy 234 GlyLeuThrIlePheTrpLeuProGluSerProArgTrpLeuValSerLysGlyArgMet 253
Db 778 GAGCTGACTGCTCTTCTCTGCTTGAATACCAAGGTGGCTTGTNACCAAGGAAGATG 837
Qy 254 AlaGluAlaLysLysValIleGluIleLeuArgGlyLysAspAspValSerGlyGluLeu 273
Db 838 GCGAGGGGGAAGAGATGTCANAGGCTGGGGGAAGAGATGTCANGGGAGANG 897
Qy 274 SerLeuLeuLeuGluGlyLeuGluValGlyIleAspThrSerIleGluGluTrpIleIle 293
Db 898 GCTCTTCTAGTTGAAGTTTGGGGTGGTAAATACACGTAATTNAGATACATCTT 957
Qy 294 GlyProAlaThrGluAlaIleAspAspLeuValIleThrAspGlyAspLysGluGlnIleThr 313
Db 958 GGAACCTGCAACCGAGGAGCGCATGATCTGTAACTGACGCTGTATAGGAACAATCA 1017
Qy 314 LeuTrpGlyProGluGluGluValGlnSerTrpIleAlaArgProSerLysGlyProIleMet 333
Db 1018 CTTATGGGCTGTAAGAGGCGCATGATGATCTGCACTTCTTAAAGGACCATCATCTG 1077
Qy 334 LeuGlySerValLeuSerLeuAlaSerArgIleGlySerMetValaGlnSerValPro 353

```

```

Db 1078 CTGGAGAGTGGCTTCTCTTCATCGTCATGAGGACATGGTGAACGAGTGTACCC 1137
Qy 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
Db 1138 CTTATGATCCGATTGTGACACTTTTGTGTAGTGTCCATGAGATATATCTCCACCTGGA 1197
Qy 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
Db 1198 GGAAGTATGAGAGACACTTGTCTCAAACTTTGAGATATGTTCACTGATCAGATCAG 1257
Qy 394 HisAlaIysAsnGluGlnTTPAspGluGluAsnLeuHisArgAspAspGluTyrAla 413
Db 1358 CATGCAAAATATGACGATGGATGAGATGAGAAATCTTCAAGGATGAGAGAGATACCA 1317
Qy 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
Db 1318 TCTGATGGTGAAGGAGTGAATGAGACATCTCCATGACCCATTCCTGTCAGGAG 1377
Qy 434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSer 453
Db 1378 GCAACAGGTGGGAGAGGAGAGACATTTGACACCATGTCACCCGTGGAAGTGTGAGC 1437
Qy 454 MetArgArgGlnThrLeuLeuGlyGlyGlyGlyAspGlyValSerSerThrAspIleGly 473
Db 1438 ATGAGAAAGGCAAGGCTCTTAGGGGAGGTGAGATGATGATGAGCAGCATGATTCGGT 1497
Qy 474 GlyGlyTTPGlnLeuAlaTTPLeuTTPSerGluGlyGlyGlyGlyAsnGlyArgGlyGlu 493
Db 1498 GGGGATGGACAGCTTGTGAAATGTCAGAGAAAGGAGGAGGATGTAAGAAAGGAA 1557
Qy 494 GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer 513
Db 1558 GGTGGTTTCAAAAGAGTCACTGACCAAGAGGAGTTCCTGCTCAAGAGGGGCTCA 1617
Qy 514 IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGlnPheValHisAla 533
Db 1618 ATTGTTTCACTTCCGAGTGGTGGCATGTTCTTGGAGTGAAGTGTGATCATCTCTCT 1677
Qy 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLysAlaGluProArgMetSerAsp 553
Db 1678 GCTTATGATGACATGACACTTTTCTCAAGGGTCTTGGAACCAAGCATGTCACAT 1737
Qy 554 AlaAlaMetValHisProSerGlnValAlaAlaIleArgLysSerArgTyrLysAspLeuPhe 573
Db 1738 GCTGCCATGCTTCAACCATCTGAGGTACTGCCAAGGTCACGTTGGAAGATTTGTT 1797
Qy 574 GluProGlyValArgArgAlaLeuLeuValGlyGlyIleGlnIleLeuGlnPhe 593
Db 1798 GAACCTGAGATGAGGCGTGCCTGTAGTGGTGTGAATTCAGATCTTCAACAGTTT 1857
Qy 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla 613
Db 1858 GCTGGAATTAACGGGTCTCTACTATACCCCAAAATCTTGACCAAGCTGTGTGCA 1917
Qy 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
Db 1918 GTTATCTCTTCCAAATTTGGTCTCAGCTCGGATACGATCCATCTTATCATGTTCTCTC 1977
Qy 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
Db 1978 ACTACTTACTATATCTCTCTTGCATTTGCTTGCATGCTGCTTATGATCTTCCGGA 2037
Qy 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
Db 2038 AGAAGGTTTTTGGCTGAGGCAAAATTCATTCATGATGACATCTCTGATATCTGAT 2097
Qy 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
Db 2098 GTGTCCATCTATATGATTTGGGTACACTAGCCCATGCTTGTCTTCCACATCATGCTT 2157
Qy 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
Db 2158 ATGCTCTACTCTGCTGCTTCTGTATGGATTTGGTCCATCCCAAACTTTATGTGCA 2217

```

```

Qy 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTyr 733
Db 2218 GAGATCTTCCAAACAGGTTTCGTGCGCTCTGATTTCCATTTGCTTACATTTCTCG 2277
Qy 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
Db 2278 ATCCGAGATATCATGTCACCTACCTGCTTCTGTATGCTGAATGCTATGAGACTGGCG 2337
Qy 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773
Db 2338 GGTGTTTTCAGACATATGACGTGTATGCTGTGATTTCTTGTGCTCTTCAAG 2397
Qy 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValAla 793
Db 2398 GTCCCTGACACAAAGGGAGTCCCTTGAAGTTTACCGAATCTTTCGATGTCG 2457
Qy 794 LysGlnAlaAlaAlaLysAla 800
Db 2458 AAGCAAGCGGCTGCCAAAGCC 2478

RESULT 5
US-10-051-902-7
; Sequence 7, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hite, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: Glycine max
US-10-051-902-7

Alignment Scores:
Pred. No.: 5,53e-274 Length: 2601
Score: 2674.00 Matches: 522
Percent Similarity: 82.20% Conservative: 92
Best Local Similarity: 69.88% Mismatches: 115
Query Match: 65.62% Indels: 18
DB: 14 Gaps: 8

US-10-051-909-32 (1-800) x US-10-051-902-7 (1-2601)
Qy 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr 75
Db 175 ATGAAAGTGGCTCTCTTGTGCTATTTGCCCTTCCATTTGATATTTCTTCAAGATGG 234
Qy 76 AspAsnAlaThrIleAlaAlaAlaValLeuTyrIleLysGluPheGlnLeuGlnAsn 95
Db 235 GATATGCTATCATCGCCGGGCTAATGATTAATTAAGAAAGACCTTGTGGAAACA 294
Qy 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db 295 -----ACTATGAAAGGCTTGTGTGGCATGTCCTGATTTGAGCAACGATATACCC 348
Qy 116 ThrPheSerGlyProLeuSerAspSerIleGlyValArgArgPrometLeuIleLeuSerSer 135
Db 349 ACAAGCTCTGATCTTATAGCGGATTTGGCTGCTGCGGACCCACATGATATATCTCATCT 408
Qy 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTTPSerProAsnValTyrValLeuLeu 155
Db 409 GTGCTCTATTTCTTGGGAGTTTGGTATGCTGTGATGCCCAATGATGATGTGTC 468

```

156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
 469 TTGGCAGAGCTACTTATGATGATTTGGGATTTGGCTTGTCTGATCTTGTCCCGGCTAT 528
 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195
 529 ATATCTGAACCGGCGCTCTGTAATAAGGGGGTCTGTAATACGCTTCTCATGTTCACT 588
 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
 589 GGCTCTGGAGGAAATGTTTGTGTAATGATGTTTGGCATGTGATGATGATGATGATGATG 648
 216 ProAspThrArgIleLeuLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
 649 CTTAGCTGAGGCTATGCTGGGGTCTGCTCATTTCTCTCTCTGTAATTTTGATTTG 708
 236 ThrIlePheTyrLeuProGlnSerProArgTyrPheValSerIleGlyArgMetAlaGlu 255
 709 ACCATTTTCTTCTGCGCCGAGTCTCTCGGTGGCTGTCTCAGCAAGAGAGATGCTGAG 768
 256 AlaIleValValLeuGluIleLeuArgGlyIleAspAspValSerGlyGluLeuSerLeu 275
 769 GCTAAGAGAGTCTCCAAAGATTCGCGAGAGAGAGATGTCTCAGGCGAGATGCGATTTG 828
 276 LeuLeuGluGlyLeuGluIleValGlyIleAspThrSerIleGlyIleTyrIleIleGlyPro 295
 829 CTGCTTGAAGTCTCTCGGATTTGGGGGTGATCATCTATGAAAGATGATGATTTGGCCCT 888
 296 AlaIleGluAlaIleAspAspLeuValThrAspGlyAspIleGluGlnIleThrLeuTyr 315
 889 GCTGACAGATGTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 948
 316 GlyProGluGluGlyGlySerThrIleAlaArgProSerIleGlyProIleLeuGly 335
 949 GGATCCCAAGCAGGCTTCTTGTGTTATCAAAACCTGTCTGAGCAGAGTTCTATTTGCG 1008
 336 SerValIleSerLeuAlaSerArgIleGlySerMetValAsnGlySerValProLeuMet 355
 1009 -----CTTGGCTCAGCAGCAGTGAAGCATCATCAACCAAGCATGCGCTCATG 1056
 356 AspProIleValThrLeuPheGlySerValIleGluAsnMetProGlnAlaGly 373
 1057 GATCTCTGGTGAACCTGTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1116
 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
 1117 GGAACCATGCGAAGCATCTGTTCCAAATTTTGAAGCATGTTCAGCATGCTGACGCG 1176
 394 HisAlaIleAsnGluIleThrAspGluGluAsnLeuHisArgAspAspGluIleTyrAla 413
 1177 CATGCTAAATTTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1236
 414 SerAspGlyValGlyIleAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
 1237 TCAGATGCAACCGTGGGAGCTCCGATGATGATGATGATGATGATGATGATGATGATG 1296
 434 AlaThrGlyAlaGluGlyIleAspIleValIleHisIleGlyHisArgGlySerAlaLeu 452
 1297 ACAACAGCTTGA-----AAGACTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1353
 453 SerMetArgArgGluThrLeuLeuGlyGlyIleGlyIleGlyValSerSerThrAsp 471
 1354 AGCATGAGGCGTCAACAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1413
 472 IleGlyGlyIleTyrGluLeuAlaTyrIleAspSerGlyIleGlyGlyIleGlyAsnGlyArg 491
 1414 ATTTGTTGGTGGCTGCACTGCGATGAGAAATGAGATGATGATGATGATGATGATGATG 1470
 492 LysGluGlyIlePheIleArgValTyrLeuHisGluGluGlyValProGlySerArgArg 511
 1471 CAACAAGAGGGTTTAAAGATTTATTAATCAATGAGAGGGGATTTCTGATCTCTGCTGT 1530

512 GlySerIleValSerLeuProGlyGlyIleAspValPheGluGlySerGluPheValHis 531
 1531 GGATCCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1572
 532 AlaAlaIleValSerGlnSerAlaLeuPheSerIleGlyLeuAlaGluProArgMet 551
 1573 GCTGCTGCTGTTGTAAGCAACCGCTCTTATCTCAAGAGCTTATGATGATGATGATGATG 1632
 552 SerAspAlaIleMetValHisProSerGluValAlaIleGlySerArgTyrPheAsp 571
 1633 GTTGGGCTGCAAAATGTTTACCCATCTGAGCAGCTTCAAGAGGCGCAAGTTGAAAGCT 1692
 572 LeuPheGluProGlyValArgArgAlaLeuValGlyValGlyIleGlnIleLeuGln 591
 1693 CTTCTTGAACAGGAGGTTAACATGATGATGATGATGATGATGATGATGATGATGATGATG 1752
 592 GlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluAlaGly 611
 1753 CAGTTTCAAGGATTAATGAGGGTCTTATTAACACACTCAAACTCTTGAAGAGCGCGT 1812
 612 ValAlaValIleLeuSerIlePheGlyLeuSerSerAlaSerIleLeuIleSer 631
 1813 GTTGAAGTTCTTCTTCAATATGAGATGATGATGATGATGATGATGATGATGATGATGATG 1872
 632 SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuMetAspLeu 651
 1873 GCTTTCACACCTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1932
 652 SerGlyArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle 671
 1933 TCAGCAGAGAGCGAGTGTCTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 1992
 672 LeuValIleSerAlaLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal 691
 1993 TTGGTCATGGAACCTGATTAATTTTGGCAATGTGCGCATGAGCAATCTCAACGTA 2052
 692 SerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu 711
 2053 TGCGTGTGTTATTTCTGCTGCTTGTATGATGATGATGATGATGATGATGATGATGATG 2112
 712 CysAlaGluIlePheProThrArgValArgGlyLeuCysAlaIleCysAlaPheThr 731
 2113 TGCTCAGAGATTTTCCCACTAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2172
 732 PheThrIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGly 751
 2173 TTCTGATGAGCATCATCATCATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2232
 752 LeuAlaGlyValPheSerIleTyrAlaValAlaCysLeuIleSerPheValPheValPhe 771
 2233 CTGTGTGTGATTTCCGATTTAGCAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2292
 772 LeuIleValProGluThrIleGlyMetProLeuGluValIleThrGluPhePheAlaVal 791
 2293 TTGAAGGTTCCAGAAACAAAGGCGATGCCCTTGAATCATCTCGAATCTTTCGTGT 2352
 792 GluValIleGlnAlaAlaAla 798
 2353 GGAGCAAGCAGGCTGCTTCT 2373

RESULT 6
 US-10-051-909-7
 ; Sequence 7, Application US/10051909
 ; Publication No. US20020199217A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: HelencJarvis, Tim
 ; APPLICANT: Hiltz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: B81163 US CIP
 ; CURRENT APPLICATION NUMBER: US/10/051,909

```

:
:
: CURRENT FILING DATE: 2002-01-17
: PRIOR APPLICATION NUMBER: 60/0083,044
:
: PRIOR FILING DATE: April 24, 1998
:
: NUMBER OF SEQ ID NOS: 38
:
: SOFTWARE: Microsoft Office 97
:
: SEQ ID NO 7
:
: LENGTH: 2601
:
: TYPE: DNA
:
: ORGANISM: Glycine max
:
: US-10-051-909-7

```

Alignment Scores:	
Pred. No.:	5.53e-274
Score:	2674.00
Percent Similarity:	82.20%
Best Local Similarity:	69.88%
Query Match:	65.62%
DB:	14
Gaps:	8

US-10-051-909-32 (1-800) x US-10-051-909-7 (1-2601)

QY MetSerGlyValAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTP 75
 Db 175 ATGAAGATGCGGCTCTTGTGCTATTGTGCCCTCCATTGGTAATTTCTCCAGAGATGG 234
 QY AspAsnAlaThrIleAlaIleAlaValLeuTyrIleLysGlnIupheGlnLeuGlnAsn 95
 Db 225 GATATGTCTACCATCGCGGGCTAATGTGTTACATTAAAGAAACCTTGCTTGGGAACA 294
 QY 96 GluProThrValGlnGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
 Db 295 -----ACATATGAAAGGCTTGTGGTGGGCATGTCCCTGATGTGACCAACGGTAATCACC 348
 QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyAsnArgProMetLeuIleLeuSerSer 135
 Db 349 ACATGCTCTGGTCTCTAATACGGAGATTGGCTGGCGGACCCATGATGATATCTACTT 408
 QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnAlaTyrValLeuLeu 155
 Db 409 GTGCTCTAATTTCTTGGGTGTGTTGGTGAATGCTGTGATCCCAATATGATGTGTGTC 468
 QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaIleThrLeuValProLeuTyr 175
 Db 469 TTGGCGAGGCTCACTTGATGGAATTTGGGATTTGGCTTGCTGTGACTCTTGGCCGGCTCAT 528
 QY 176 IleSerGlnIleAlaProSerGlnIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195
 Db 529 AATATGTGAAACGGCGCGCTGTAATAAGGGGGTGGTGTGAATACGCTTCTCGAGTTCAGT 588
 QY 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
 Db 589 GGCTCTGAGAGAAATGTTTGTGTGCTGATGTAATGTTTGGCATGTGATGATCCCGCG 648
 QY 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
 Db 649 CCTAGCTGAGGCTCATGCTTGGGGTGTGTGCTAATCTCTCTCTTGATTTTGGCATTTG 708
 QY 236 ThrIlePheTyrLeuProGlnSerProArgTrpLeuValSerIleGlyAsnMetAlaGln 255
 Db 709 ACCATTTTCTTTCGCGCCAGATCTCTCGGTGGCTGGTACGCAAGAGAAAGATGCTCGAG 768
 QY 256 AlaIleLysValLeuGlnIleLysLeuAcArgGlyLysAspAspValSerGlyGlnLeuSerLeu 275
 Db 769 GCTTAAGAAAGTGTCTCCAAAGATTGCCCGAAGGAGAGATGTGTACGGCGAGATGGATTG 828
 QY 276 LeuLeuGlnGlyLeuGlnIleValGlyIleAspThrSerIleGlnGlnTyrIleIleGlyPro 295
 Db 829 CTGGTTGAAGTCTCCGGATTTGGGGGTGATACATCTATCGAAGAGTCAATAATGGCCCT 888
 QY 296 AlaThrGlnAlaIleAspAspLeuValThrAspGlyAspIleGlnGlnIleThrLeuTyr 315
 Db 889 GCTGACGATGTGCTAATGTCTATGAACATGCAACAGAAAGAAATAAATTCATATAT 948

QY	316	GIYProgluGIYgluYInserTrrIlaIArgProserIySGIYProIleMeIeUglY	315	:::	
Db	949	GGATCCCAAGAGGCGCTTTCTGGTTATCAAAACCTGTCACTGACAGAGTTCATTGGC	1008	:::	
QY	336	SeValIeuSerIeuIaIaSerArGhIscGIYSerMetValaEngInSerValProIeuMet	355		
QY	1009	-----CTTGGGTGCAACCATGGAAGCATATCAACCAAGACCTCCCTCATG	1056		
QY	356	AspProIleValThrLeuPheGlySerValHisGluasnMetProGlnaIagly-----	373		
Db	1057	GATCCTCGTGAGCACTGTTTGTATGATCATAGAAAGCTCCCGAGACGAGAGCAGA	1116		
QY	374	GlySerMetArgSerThrIleuPheProaspPheGlySerMetPheSerValThrAspGln	393		
Db	1117	GGAAAGCATGCGAAGCACTGTCTTCCAAATTTGAAAGCAATGTTGACGACTGCTGACCG	1176		
QY	394	HisAlaIalysaEngIuGInTrpAspGluGluasnLeuHisArgAspAspGluGluTrpAla	413		
Db	1177	CATGCTAAATTTGAACAATGGGATGAAGAAAGCTTACAAAGGAACGTGAAGACTACATG	1236		
QY	414	SerAspGlyAlaGlyIyAspTyrgIuaspAsnLeuHisSerProIeuLeuSerArGln	433		
Db	1237	TCAAGATGCAACCCGTGGGAGCTCCGAGATATTTGGACAGTCCCTTATCTCAGCCAA	1296		
QY	434	AlaThrGlyAlaGluGluIyAspAlIeValHisGlyHisArgGlySerAlaIeu---	452		
Db	1297	ACAACAAAGCCTTGA--AAAGACTTACTCTCTCTCTCTCCACATGGAGATGCTTGGC	1355		
QY	453	SerMetArgArGInThrIleuLeuGlyIyGluIy--GlyAspGlyValSerSerThrAsp	471		
Db	1354	AGCATGAGGCTCACAGTACATCTATGCAAGAGGTGAGTGCACAAAGTGTAGTACAGT	1413		
QY	472	IleGlyIyGlyTrpGlnLeuAlaTrpIySerGluIySGluIyGluasnGlyArg	491		
Db	1414	ATTGGTGTGCTGGCAACTGGCATGGAATAGCACTGATAA--GGTGAAGATGGAATA	1470		
QY	492	LysGluGlyGlyPheIyAsrValTyrlleuHisGlnGluGlyValProGlySerArGrG	511		
Db	1471	CAACAGAGAGGGATTTAAAGGATTTATTTACATGAGGAGGAGTTCTGCATCTCGTGT	1530		
QY	512	GlySerIleValSerIeuProGlyIyGlyAspValPheGluGlySerGluPheValHis	531		
Db	1531	GGATTCATTGTATCATTCATTCCTCCGGTGAAGC-----GAATTGTCCAG	1572		
QY	532	AlaAlaIalaleuValSerGlnSerIalaleuPheSerIySGlyLeuAlaGluProArGmet	551		
Db	1573	GCTGTGCTCTGTATGAGCCAAACCGCTCTTACTCCAAAGAGCTTATTTAGTGAACCCA	1633		
QY	552	SerAspAlaIaIeMetValHisProSerGluValAlaAlaIySGlySerArGrTrpIyAsp	571		
Db	1633	GTTGGGCTCGCAATGGTTCACCCATCTGAGACAGCTTCAAAAGGGCGCAAGTTGGAAACCT	1692		
QY	572	LeuPheGluProGlyValArgArgAlaIeuIeuValGlyValGlyIleGlnIleuGln	591		
Db	1693	CTTCTTGACCAAGGGGTTAAGCATGTCATGTGTGTGAAGTTGGAATVCAAAATCTTCAG	1755		
QY	592	GlnPheAlaGlyIleasnGlyValLeuIyTrpThrProGlnIleLeuGluGlnaIagly	611		
Db	1753	CAGTTTCAGGATTAATAGGGGTTCTATATTAACAACACTCAATCTTGAAGAGGCCGGT	1812		
QY	612	ValAlaValIleLeuSerIyPheGlyLeuSerSerAlaSerAlaSerIleIeuIleSer	631		
Db	1813	GTTGAAGTCTTCTTTCAGATATAGGCACTTGGCTCAGAGTGGGATCTATTCCTTACGT	1872		
QY	632	SerIeuThrThrIleuIeuMetIeuProCysIleGlyPheAlaMetLeuIeuMetAspIeu	651		
Db	1873	GCTTTCACAACCTTCTTATGCTTCCCGTATAGGCGTGAAGCATGAAGCTCATGGATGT	1933		
QY	652	SerGlyArgArPheIleuLeuGluIyThrIleProIleIleuIleAlaSerIeuValIle	671		
Db	1933	TCAGGCAAGAGCACTTGTACTTATCTACATCCCGGTCGATGTGTGTCACTATATT	1992		
QY	672	LeuValValSerAsnLeuIleAspIeuGlyThrIleuAlaHisAlaIeuIeuSerThrVal	691		

Db 1180 CATGGAAGCTCAAGTCAGGAGCTCAAGCGGAGGAGCGGAGTATGAGGAGTTGAGCT 1239
 Qy 475 GATTPGInleuAATrPlySTrSergLyuSgInglYgluAenGlyArglySgInglY 494
 Db 1240 GGAATGGCAATGGGATGGAATGACGGAAGAGAGATGAATCGGACAGAGAGAA--- 1296
 Qy 495 G1PheLySargValYrLeuHtSgInglYgluAValProglYSerArgSgInglY 514
 Db 1297 -----GAGGTTTCCAGAGATCTCGAGCGTCTCAATT 1329
 Qy 515 ValSerLeuProglYglYglYAspValPheglYglYSerLyuPheValHisAlaAla 534
 Db 1330 GTTTCATTCCTGGTGGTGAATGAAACCGATG---GGAATTTTGTCAGACGCTCTCT 1386
 Qy 535 LeuValSerGInSerAlaLeuPheSerLySgInglYgluAgluPProArgMetSerAspAla 554
 Db 1387 TTGGTTAGCCCAACGCTCTTTATTCGAAAGCCTTTCGAAAGAACATCAATGCTCTCT 1446
 Qy 555 AlaMetValHisProSerGluValAlaAlaLySgInglYSerArgTrpLyAspLeuPheGlu 574
 Db 1447 GCTATGATACATCCATCCGAA---ACAACTAAGGTCATTTGGCATGATCTTCATGAT 1503
 Qy 575 ProglYValArgArgAlaLeuValGlyValGlyIleGInlIleuGIngluPheAla 594
 Db 1504 CCTGAGTCAAGCGGTGATAGTCTAGAGATTGAGCTTCAATCTTCAGCAGTCTCA 1563
 Qy 595 GlyIleAenGlyValLeuTyTrpProGInlIleuGIngluAglYValAlaVal 614
 Db 1564 GGCATCAACGAGGTCCTTACTACACCGCAATCTTGAAGCAGCGGCTGTCGGAGTC 1623
 Qy 615 IleuSerLySgInglYleuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThr 634
 Db 1624 CTACATGAGAACATGGGATTAAGTCTTCTCAGCATCTTACTTAAAGTCATGACA 1683
 Qy 635 ThrLeuMetLeuProCysIleGlyPheAlaMetLeuMetAspLeuSerGlyArg 654
 Db 1684 ACCTTTGATGATTAAGTCAATAGCTGTGCAATGAGGCTCATGAGATCTTCTGATGCA 1743
 Qy 655 ArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleuValVal 674
 Db 1744 AGGACCTGCTTCTACACGATACCAATCTGATGCAATCTCTATTTGGTTTAAATATC 1803
 Qy 675 SerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuSerThrValSerValIle 694
 Db 1804 TCAATCTTGTTCATACATGACGATGCGACGGGCTTTCATCAACGTAAGCTGTG 1863
 Qy 695 ValYrPheCysCysPheValMetGlyPheGlyProIleProAsnIleuGlyValGlu 714
 Db 1864 CTCTACTCTGCTCTTGGTATGGGTTTGGTCTGCTCAAAACATCTCTGTTCAAG 1923
 Qy 715 IlePheProThrArgValArgGlyLeuGlySerIleAlaIleCysAlaPheThrPheTrpIle 734
 Db 1924 ATTTTCCAACTCGAGTCGCGGAATCGCATCGCATCTGCGCATCACTTCGAGATC 1983
 Qy 735 GlyAspIleIleValThrTySerLeuProValMetLeuAsnAlaIleGlyLeuAlaGly 754
 Db 1984 TGTGATATATGATCTTACATGATCTCCCGTCTGCTCAATTCATGAGATGATGCTGT 2043
 Qy 755 ValPheSerIleTyAlaValAlaCysLeuIleSerPheValPheValPheLeuValVal 774
 Db 2044 GTGTTTGAATGTACGCAATCGTATGTCATTTGATGGGTCTTGTGTTCAATTAAGTC 2103
 Qy 775 ProglYThrLySgInglYMetProLeuGluValIleThrGluPhePheAlaValGlyAlaLys 794
 Db 2104 CCGGAACCTAAAGGATGCGACTTGAAGTCACAGAGTCTTTTGTGTTGAGACTGA 2163
 Qy 795 GlnAlaAlaAla 798
 Db 2164 CAAGCTGAAGCT 2175

; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE REFERENCE: SCRIPI300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 2254
 ; LENGTH: 2205
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; US-09-938-842A-2254
 Alignment Scores:
 Pred. No.: 1,37e-231 Length: 2205
 Score: 2275.00 Matches: 462
 Percent Similarity: 75.89% Conservative: 111
 Best Local Similarity: 61.19% Mismatches: 146
 Query Match: 55.83% Indels: 36
 DB: 10 Gaps: 14
 US-10-051-909-32 (1-800) x US-09-938-842A-2254 (1-2205)
 Qy 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuGInglYTrp 75
 Db 1 ATGAAGGAGAGGACTCTGTTGCTCTCGCCGCAATCGGCAATTTCTTCAAGATGG 60
 Qy 76 AsnAsnAlaThrIleAlaAlaValLeuValIleValSgInglYpPheGInleuGInAsn 95
 Db 61 GACATGCGACCATTCGTCGAGCTATGATTAATCAACAAAGACTGAATCTA----- 114
 Qy 96 GluProThr---ValGInglYleuIleValSerMetSerLeuIleGlyAlaThrIleVal 114
 Db 115 ---CCAACTCTGTTCAAGTCTTGTCTGCTGCTATGATGATGCGTCAACGCTATC 171
 Qy 115 ThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSer 134
 Db 172 ACGACTTCTCAGGACCGATATGATTTGGCTCGGACAGACGCCCATGCTCATTTATCA 231
 Qy 135 SerIleLeuTyrrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrrValLeu 154
 Db 232 TCAGTTATGATTTTCGTCGCGGTTTGAATATGTTGCTCTCCCAATGTCATATGTTTG 291
 Qy 155 LeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeu 174
 Db 232 TGCTTGTCTAGCTCTTAAATAGGTTTGTCGCGGCTCGGCTTCACTTGTCTCTGT 351
 Qy 175 TyrrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGInPhe 194
 Db 352 TACATTTCTGAACCGCTCTCCGAGATCAGAGGACAGTAAATATCACTCCCTCAGTTT 411
 Qy 195 SerGlySerGlyGlyMetPheLeuSerTyrrCysMetValPheGlyMetSerLeuSerPro 214
 Db 412 CTGGCTCTGATGATATTTTGTGCATACGTATGTTTTCATATATCCCTGATGATC 471
 Qy 215 SerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhePheGly 234
 Db 472 TCCCTTACTGAGAGACCATGCTCGGTCTCTCGATCCCTCTCTTATTTATTTGTT 531
 Qy 235 LeuThrIlePheTyrrLeuProGluSerProArgTrpLeuValSerLySgInglYArgMetAla 254
 Db 532 CTCACGGTGTATTATTTGCCCCGAGTCTCTGTTGGCTGTGTTAGTAAGAAAGATGAC 591

QY 255 GIUAlaLysValLeuGlnLysLeuAArgGlyLysAspAspValSerGlyLeuLeuSer 274
 Db 552 GAGGCTAAGCGAGCTTCTTCAACAGTTATGTGCAAGAAATGTTATCCATAGATGGCT 651
 QY 275 LeuLeuLeuGlnGlyLeuGlnValGlyLysPheThrSerLeuGlnLysLeuLeuGly 294
 Db 652 TTACAGTGTGAAGAGCTAGATATAGAGAGAAACAAATGAAAGATCTCTAGTAAGT 711
 QY 295 ProAlaThrGlnAlaAlaAspAspLeuValThrAspGlyAspLeuGlnGlnLeuThr 313
 Db 712 TTGAGGATCATGAAAGGTATGATACACTTGAAACCGTTGATGAGATGCAAAATGCGG 771
 QY 314 LeuTYrGlyProGlnLysLysLeuSerTrpIleAlaArgProSerLeuGlyProIleMet 333
 Db 772 CTTTATGGAACCCAGCAAGATCATCTGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCT 825
 QY 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
 Db 826 ---AATAGCTCACTGGGCTACGCTCTGCGCAAGGAGCTTAGCAACCAAGCATGATC 882
 QY 354 LeuMetAspProIleValThrLeuPheGlySerValHisGlnAsnMetProGlnAlaGly 373
 Db 883 CTTAAAGATCCGCTGCTCAATCTTTTGGCACTCTCCAGAGAGATGCCAGAGACAGCG 942
 QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
 Db 943 GGAACACCTCGAGAGTGGGATTTCCCTCACTTCGAGAGATTTCAAGTACTGCTCCGAT 1002
 QY 394 ---HisAlaLysAsnGlnLysPheAsp-----GlnGlnAsnLeuHisArgAsp 408
 Db 1003 GCGGCTCAAGCTAAACCGGCTCATTTGGGAAAGACATAGAGACCATTAACAAGAGC 1062
 QY 409 AspGlnGlnLysValSerAspGlyAlaGlyLysAspTYrGlnAsp-----AsnLeu 425
 Db 1063 AATGATGACTAGTCACTGATATGATGCGGATGATGATGATGATGATGATGATGATGAT 1122
 QY 426 HisSerProLeuLeuSerArgGlnAlaThrGlyAlaGlnGlyLysAspIleValHis 445
 Db 1123 CGTACCCCTTATGCTCGCGCCAGACACACAGCATGAGC---AAGATATGATCCCACT 1179
 QY 446 GlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlnGlyLysAsp 465
 Db 1180 CTTAACAAGGAGACCTTATAGCATGAGACACACAGCATGAGC---GCGAAC 1236
 QY 466 GlyValSerSerThrAspIleGlyGlyTYrGlnLeuAlaTrpLysTrpSerLeuLys 485
 Db 1237 GCGGAAAGTACGATGGAATGTGTGTGCTGGCAATGAGATATAGATACGAAACGAT 1296
 QY 486 GlnGlyLysAsnGlyArgGlyGlnGlyPheLysArgValTYrLeuHisGlnGlnGly 505
 Db 1297 GAA-----TACAAAGATATTAATCTTAAGAAAGATGGA 1329
 QY 506 ValProGlySerArgArgGlySerIleValSerLeuProGlyGlyLysAspValPheGln 525
 Db 1330 GCTGAA---TCTCGCGGTGCTCGATCATCTATCTCCGAGGAT-----GAT 1380
 QY 526 GlySerGlnPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLeuGly 545
 Db 1381 GAGGACACTACATTCACGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
 QY 546 LeuAlaGlnProArgMetSerAspAlaAlaMetValHisProSerGlnValAlaLys 565
 Db 1441 -----GTTCTATGATCCGCGCATGATCCCGCGAGAAATGCTGCTCT 1485
 QY 566 GlySerArgTYrPheAspLeuPheGlnProGlyValArgArgAlaLeuLeuValGlyAl 585
 Db 1486 GAGCAGCTCTGCTGCTCTCTTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1545
 QY 586 GlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTYrTYrThrProGln 605
 Db 1546 GGCATTCAAAATATCTGACAGATTTTCAGGTATCAATGAGGTTCTCTACTACACTCTGAG 1605

QY 606 IleLeuGlnGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSer 625
 Db 1606 ATTCTCGAAGCGGCTGGGTATGATATCTTCTTTCGAGCCCTGGAGATAGTTCATCTCT 1665
 QY 626 AlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAla 645
 Db 1666 GCGTATCTCCATCAGCGGCTTAAACAATTAATCAATGCTGCTCCAGCATATGCTGCTGCC 1725
 QY 646 MetLeuMetAspLeuSerGlyArgArgPheLeuLeuGlnTYrThrIleProIleLeu 665
 Db 1726 ATGAGACTCATGATGATCCGGAAGAGGATTAATTAATCTTCTGACATATCCAGATCTTC 1785
 QY 666 IleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHis 685
 Db 1786 ATTTGCTACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1845
 QY 686 AlaLeuLeuSerThrValSerValIleValTYrPheCysPheValMetGlyPheGly 705
 Db 1846 GCACACACTCTCCAGAGGTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1905
 QY 706 ProIleProLeuIleLeuCysAlaGlnIlePheProThrArgValArgGlyLeuCysIle 725
 Db 1906 CCGATTCGAACATCTCTGTTGAAATCTTCCCAACAGAGTCCGTTGCTGCAATC 1965
 QY 726 AlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTYrSerLeuProVal 745
 Db 1966 GCGATATGTGCTATAGTCTTCTTGAATGAGACATTAATGTCAGTACTGCTCCGCTT 2025
 QY 746 MetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTYrAlaValAlaCysLeuIle 765
 Db 2026 CTCCTCAGCTGATCGGACTGTTGTTGTTTTCAGCATTTACGCTGCGTTGGCTATTC 2085
 QY 766 SerPheValPheValIlePheLeuValProGlnThrLysGlyMetProLeuGlnValIle 785
 Db 2086 TATGATCTTCTGTTTACATTAAGTCCCGAGACTTAAGGCAAGCTTGTGAAGTTATC 2145
 QY 786 ThrGlnPhePheAlaValAlaGlyAlaGlnAlaAlaAlaLysAla 800
 Db 2146 ACAGACTACTTGTGCTTGGAGCT---CAAGCTCAAGCTTGTGCT 2187

RESULT 9
 US-10-051-902-9
 ; Sequence 9, Application US/10051902
 ; Publication No. US20020178468A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Hitz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Kinney, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BR-1163
 ; CURRENT APPLICATION NUMBER: US/10/051,902
 ; CURRENT FILING DATE: 2002-01-17
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
 ; NUMBER OF SEQ ID NOS: 30
 ; PRIORITY FILING DATE: 1999-04-14
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 9
 ; LENGTH: 1692
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; US-10-051-902-9

Alignment Scores:
 Pred. No.: 2,27e-167 Length: 1692
 Score: 1670.50 Matches: 333
 Percent Similarity: 79.75% Conservative: 57
 Best Local Similarity: 68.10% Mismatches: 84
 Query Match: 40.99% Indels: 15
 Db: 14 Gaps: 8

US-10-051-909-32 (1-800) x US-10-051-902-9 (1-1692)


```

Qy 308 AspylsgluniliethrleuTyrlProgluglunlglnsertPrleAlaArgPro 327
Db 21 GAAAAAGTCAAAATTAAGTGTATGACACAGAACAGCCAGTCTGGGTGTACAGACT 80
Qy 328 SerlysglyProilemetleuGlySerValleuSerleuAlaSerArgHisGlySerMet 347
Db 81 GTTGCTGACCAAAATTCCTGTGGC-----CTTGATCTAGAGAAAGGAAGCAATG 128
Qy 348 ValanglnsertValProleuMetAspProileValThrleuPheGlySerValHisGlu 367
Db 129 GCAAAATCCAGC---AGCTAGTGAGACCCCTTAGACCCCTTTGGTAGTGTACATAG 185
Qy 368 AsnMetProglunAlaGlyGlySerMetArgSerThrleuPheProAsnPheGlySerMet 387
Db 186 AAGCTCCAGAAACAGG-----AGCACCTTTTCCACACTTTGGAGATAG 233
Qy 388 PheSerValThrAspGlnHisAlaLysanglnGlnThrAspGlnGluAsnLeuHisArg 407
Db 234 TTCAGTGTGGGGGAAATCAGCCAGAGATGAGATGGAGTGAAGAAAGCCTAGCCAGA 293
Qy 408 AspAspGlnGluTyrrAlaSerAspGlyValaglyGlyAspTyrrGluAspAsnLeuHisSer 427
Db 294 GAGGTGTAGATATGCTCTGAT-----GCTGTATCTGTAGACAAATTTGAGAGT 347
Qy 428 ProleuSerArgGlnAlaThrGlyValaglnGlyLysAspLeuAlaHisGlyHis 447
Db 348 CCATTGATCTACCGTCAACACAGAGTCTGAT---AGGACATACCTCCATCCCAT 404
Qy 448 ArgGlySerAlaLeuSerMetArgArgGlnThrleuLeu---GlyGlnGlyGlyAspGly 466
Db 405 AGTAACTCTTGA---AGCATGAGGCAAGTAGTCTTTATACATGAAATTCAGAGAACCC 461
Qy 467 ValSerSerThrAspLeuGlyGlyGlyTyrrProGlnLeuAlaThrLysTrpSerGlnLysGlu 486
Db 462 ACTGTATACTCGGATGTGGTGTGGTGGCAGTACAGATGGAATGTCTGAAAGAG 521
Qy 487 GlyGlnanglnYargLysGlnGlyGlyLysPheLysArgValTyrrLeuHisGlnGlnGlyVal 506
Db 522 GGGCCAGATGAAAGAAAGAGAGTGGCTTCAAGAGAAATATTTACACCAAGATGTGGT 581
Qy 507 ProGlySerArgArgGlySerLeuValSerleuProGlyGlyLysAspValPheGlnGly 526
Db 582 TCTGATCTAGACGCGGTGTGTGTGTTCACTCCCT---GGGGGTGATTTACCACTGAC 638
Qy 527 SerGlnPheValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeu 546
Db 639 AGTGAAGTTGTACAGGCTGCTGCTGTGAGTACGCTGCCCTTATATATAGAGACTT 698
Qy 547 AlaGlnProArgMetSerAspAlaAlaMetValHisProSerGlnValAlaAlaLysGly 566
Db 639 ATGCCTCAACGGCCAGTTCGACCACTATGATTCCTCTGAAACAAATTCAGAAAGGG 758
Qy 567 SerArgTrpLysAspLeuPheGlnProGlyValArgArgAlaLeuLeuValGlyValGly 586
Db 759 CCAAGTGTGAGATCTTTTGAACCTGGGGTGAAGACATGATGATTTGGGGGTGGGA 818
Qy 587 IleglnleuGlnGlnPheAlaGlyLysleanglyValLeuTyrrTyrrThrProGlnHis 606
Db 819 ATGCAAAATCTTACAGAGTCTCTGTATTAATGGGGTCTCTACTATACGCTCAAAAT 878
Qy 607 LeuGlnGlnAlaGlyValAlaValAlaLeuSerLysPheGlyLeuSerSerAlaSerAla 626
Db 879 CTGAGCAGGCAAGGCTTGTGTTATCTCTTTCAAGCCTAGGCTGTGTTCTTCTTCA 938
Qy 627 SerleuLeuLysSerSerleuThrThrleuLeuMetLeuProCysIleglyPheAlaMet 646
Db 939 TCTTCTTATTATAGCGGTGACCAACTGTGATGCTTCTTGTATACCATTCGACATG 998
Qy 647 LeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThrIleProleuLeu 666
Db 999 AGGCTCATGTATATTTACAGAGAGCATTTGCTGTAGTACAAATCCCTCCCTTAATA 1058
Qy 667 AlaSerLeuValIleLeuValValSerAlaLeuLysLeuGlyThrleuAlaHisAla 686

```

```

Db 1059 GCACCTCTTTCATATATAGTCTCGGAAGTCTTGATTTGGGATCCACTGCAGAAATGCA 1118
Qy 687 LeuLeuSerThrValSerValIleValTyrrPheCysPheValMetGlyPheGlyPro 706
Db 1119 TCAATCTCAACCATTAAGTATGATGCTATATTTGTTGTTCTTGTTCATAGGAGTGGACCA 1178
Qy 707 IleProAsnIleLeuCysAlaGlnIlePheProThrArgValArgGlyLeuCysAla 726
Db 1179 ATTCTTAATATACCTTTGTGAGAGATCTTCCCACTGAGTCTGTGCTGTGATTCCT 1238
Qy 727 IleCysAlaPheThrPheTrpIleglyAspIleleValThrTyrrSerleuProValMet 746
Db 1239 ATTGTGCTTACTTACTTTGGATCTGATATCATTTGATGACCTACACACTCCAGTATAG 1298
Qy 747 LeuAsnAlaIleglyLeuAlaGlyValPheSerIleTyrrAlaValValCysLeuLysSer 766
Db 1289 CTCAATCTGTAGGCTCGCTGCTGTGTTGTATTTATGCTGTGCTGTGATAGCA 1358
Qy 767 PheValPheValPheLeuLysValProGlnThrLysGlyMetProLeuGlnValIleThr 786
Db 1359 TGGGTGTTGCTCTTTTGAAGTTCCAGAAACCAAGGAGATGCCACTGGAAGTATCAT 1418
Qy 787 GluPhePheAlaValGlyAlaLysGln 795
Db 1419 GAGTCTCTCTGTCGAGCAAAACAG 1445

```

RESULT 10

```

US-10-051-909-9
; Sequence 9, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tiney, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: Bn163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 1692
; TYPE: DNA
; ORGANISM: Glycine max
US-10-051-909-9

```

Alignment Scores:

```

Pred. No.: 2,27e-167 Length: 1692
Score: 1670.50 Matches: 333
Percent Similarity: 79.75% Conservative: 57
Best Local Similarity: 68.10% Mismatches: 84
Query Match: 40.99% Indels: 15
DB: 14 Gaps: 8

```

```

US-10-051-909-32 (1-800) x US-10-051-909-9 (1-1692)

```

```

Qy 308 AspylsgluniliethrleuTyrlProgluglunlglnsertPrleAlaArgPro 327
Db 21 GAAAAAGTCAAAATTAAGTGTATGACACAGAACAGCCAGTCTGGGTGTACAGACT 80
Qy 328 SerlysglyProilemetleuGlySerValleuSerleuAlaSerArgHisGlySerMet 347
Db 81 GTTGCTGACCAAAATTCCTGTGGC-----CTTGATCTAGAGAAAGGAAGCAATG 128
Qy 348 ValanglnsertValProleuMetAspProileValThrleuPheGlySerValHisGlu 367
Db 129 GCAAAATCCAGC---AGCTAGTGAGACCCCTTAGACCCCTTTGGTAGTGTACATAG 185

```

```

Qy 368 AsnMetProGlnAlaGlyValSerMetArgSerThrLeuPheProAsnPheGlySerMet 387
Db 186 AAGCTCCAGAAACAGGA-----AGCACCTTTTCCACACTTGGAGTATG 233
Qy 388 PheSerValThrAspGlnHisAlaValSaangluGlnThrPheGluGluAsnLeuHisArg 407
Db 234 TTCATGTTGGGGAATTCAGCCAGAGATGAAGATGGGATGAGAGAAAGCTTACCCAGA 293
Qy 408 AspAspGluGluValAlaSerAspGlyAlaGlyValAspThrGluAspAsnLeuHisSer 427
Db 294 GAGGATGATGATTAATCTCTCAT-----GCTGGATGATTCGATGACAAATTTGACAGAT 347
Qy 428 ProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyValAspHisValHisGlyHis 447
Db 348 CCATTGATCTCAGCTCAACCAACGAGTCGAT---AGGACATACCTCTCATGCCCAT 404
Qy 448 ArgGlySerAlaLeuSerMetArgGlnThrLeu---GlyGluGlyValAspGly 466
Db 405 AGTAACCTTGA---AGCATGAGCGAAGGATGCTTTTACATGAAATTCAGAGAACCC 461
Qy 467 ValSerSerThrAspIleGlyValGlyValTrpGlnLeuAlaTrpValTrpSerGluVal 486
Db 462 ACTGTAAGTACTGGAGTTGGTGGTGGTGGAGCTGACGTGACATGGAATGGCTGAAGAAGAG 521
Qy 487 GlyValSaanglyArgGlyGluGlyValPheValArgValTyrLeuHisGlnGluGlyVal 506
Db 522 GGGCCAGATGGAAGAAGAGAGAGAGTGGCTTCAAGAGATATATTACACCAAGATGGTGT 581
Qy 507 ProGlySerArgArgGlySerIleValSerLeuProGlyGlyValAspValPheGluGly 526
Db 582 TCTGATCTAGACGTGGGTCTGGTTCACTCCCT---GGGCGATATTACCAACTGAC 638
Qy 527 SerGluPheValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSerGlyVal 546
Db 639 AGTGAAGCTTGTACAGCTGCTGCTGAGTACAGCTGCTGCTTATATATAGAACCTT 698
Qy 547 AlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaValGly 566
Db 699 ATGCCTCAACGGCCAGTTGGACCCAGCTATGATTCACTCTGAAACAAATTGCCAAAGGG 758
Qy 567 SerArgTrpValAspLeuPheGluProGlyValArgValAlaLeuValGlyValGly 586
Db 759 CCAGCTTGAGTACCTTTTGAACCTGGGGAAGCATGATGATGCTGGGGGTGGGA 818
Qy 587 IleGlnIleLeuGlnIlePheAlaGlyIleAsnGlyValLeuTyrThrProGlnIle 606
Db 819 ATGCAATTTCTCACAGATCTCTGTATAAATGGGCTCTTACTATACGCTCAAAAT 878
Qy 607 LeuGlnGlnAlaGlyValAlaValIleLeuSerTyrPheGlyLeuSerSerAlaSerAla 626
Db 879 CTGAGCAGAGGAGGTGGTATCTCTTCAAGCTAGGCTTGGTCTTACTCTTCA 938
Qy 627 SerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMet 646
Db 939 TCTCTTCTTATTAAGCCGTGACCAACTGTTGATCTCTCTGTATAGCCATTGCCATG 998
Qy 647 LeuLeuMetAspLeuSerGlyValArgPheLeuLeuLeuGlyThrIleProIleLeuIle 666
Db 999 AGGCTCAGAGATATTTCAGGCAAGGACTTGTGCTGACATGACATCCCTCTTAATA 1058
Qy 667 AlaSerLeuValIleLeuValValSerLeuLeuIleAspLeuGlyThrLeuAlaHisAla 686
Db 1059 GCAGCTCTTCTCATTTAGTCTCTGGGAAGTCTTGAGATTTGGGATCCACTGCAATGCA 1118
Qy 687 LeuLeuSerThrValSerValIleValTyrPheCysPheValMetGlyPheGlyPro 706
Db 1119 TCAATCTAACCATTAAGTGTATTTGCTATTTCTGTTCTTGTATGAGGATTTGACCA 1178
Qy 707 IleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuValIleAla 726
Db 1179 ATTCTTAATATACCTTTGTCAGAGATCTTCCACCTGAGTGTGGTCTCTGCAATGCT 1238
Qy 727 IleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMet 746

```

```

Db 1229 ATTTGTGCCCTTACTCTTTGGATCTGTCATATTCATTGACCTACACCTCCAGTTATG 1298
Qy 747 LeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValAlaCysLeuIleSer 766
Db 1299 CTCAATCTTGAAGCCTCTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1358
Qy 767 PheValPheValPheLeuValProGluThrArgGlyMetProLeuGluValIleThr 786
Db 1359 TGGGTGTTGCTCTTTTGAAGTTCCAGAAACCAAGGAGATGCCACTGGAAGTATCAT 1418
Qy 787 GluPhePheAlaValAlaValGlyValGlyVal 795
Db 1419 GAGTCTCTCTGTCGAGCAAAACAG 1445

RESULT 11
US-10-051-902-13
; Sequence 13, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hite, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tinney, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1487
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-051-902-13

Alignment Scores:
Pred. No.: 3,366-141 Length: 1487
Score: 1424.00 Matches: 281
Percent Similarity: 89.97% Conservative: 24
Best Local Similarity: 82.89% Mismatches: 32
Query Match: 34.94% Indels: 2
DB: 14 Gaps: 2

US-10-051-909-32 (1-800) x US-10-051-902-13 (1-1487)
Qy 462 GluGlyGlyAspGlyValAspSerThrAspIleGlyGlyValTrpGlnLeuAlaTrpVal 481
Db 12 GAGGGTGGGAGGAGCATGACAGCATGCTGTATGTTGGGGGGTGGCAACTGCGATGAAA 71
Qy 482 TrpSerGluValGluGlyValSaanglyArgGlyGluGlyValPheValArgValTyrLeu 501
Db 72 TGTGCGAGCAGCAAGGCGAGAGATGCGCAAGAGAGAGGCTTCAAAAGATCTACTTG 131
Qy 502 HisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGly 521
Db 132 CACCAAGGGGGGTGGCCGACTCAAGAGGGGCTCTGTGTTTCACTTCTGGTGGGGGT 191
Qy 522 AspValPheGlu---GlySerGluPheValHisAlaAlaAlaLeuValSerGlnSerAla 540
Db 192 GATGCCAGCAGAGGGGCGAGTGGGTTTATACATGCTGCTTGTGGTGAAGCCTCGGCT 251
Qy 541 LeuPheSerTyrGlyLeuAlaGluProArgMet---SerAspAlaAlaMetValHisPro 559
Db 252 CTTAATCTCAAGATCTTATGAAAGCGTATGCGCGGCTCCAGCCATGATTCATCA 311
Qy 560 SerGluValAlaAlaValGlySerArgTrpValAspLeuPheGluProGlyValArgArg 579
Db 312 TTGAAGGAGCTCCCAAGGTTCAATCGAAGAAATCTGTTGAACCTGAGTGAAGGCT 371
Qy 580 AlaLeuLeuValGlyValGlyIleGlnIleLeuGlnIlePheAlaGlyIleAsnGlyVal 599

```

```

Db      372 GCATTGTCGTCGGTGTGGAATTCAGATGCTTGAGCAAGTTGCTGGAATAATGAGATT 431
Qy      600 LeuTYrTYrProGlnIleuGlnIleuGlnIleuGlnIleuSerIleuPhe 619
Db      432 CTCTACTACTCTCTCAATTCCTGAGCAAGCTGCTGCTGCTCTCTCTCTCTCTCT 491
Qy      620 GlyLeuSerSerAlaSerAlaSerIleuIleuIleuSerLeuThrThrLeuMetLeu 639
Db      492 GGCCTCAGTTCAGATCAGATCAGATCAGATCAGATCAGATCAGATCAGATCAGATC 551
Qy      640 ProCysIleGlyPheAlaMetLeuMetLeuMetLeuMetLeuMetLeuMetLeu 659
Db      552 CCAAGCATTTGGTGTGAGCATGAGCATTTGATGATATCTGGAAGAAGTTTCTGTAAG 611
Qy      660 GlyThrIleProIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 679
Db      612 GGCACAATTTCCATCTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 671
Qy      680 LeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValIleValIle 699
Db      672 TTGAGTAGGAGGCCCCCAGCTGTGCTCTCCACAGTAGAGTGTATGCTACTCTGCTGC 731
Qy      700 PheValMetGlyPheGlyProIleProAsnIleuGlyValAlaGluIlePheProThrArg 719
Db      732 TTGTGCATGGGCTTTGGCCCATCCCAACATTCATGTGAGAGATTTCCCAACAGCA 791
Qy      720 ValArgGlyLeuGlyAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleVal 739
Db      792 GTCCGTGTGTCTGTGATGCTGATTTGGCCCTCCACATTCGTGATTTGAGATTTATGTT 851
Qy      740 ThrTYrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleIle 759
Db      852 ACCTACACCTGCGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 911
Qy      760 AlaValValCysLeuIleSerPheValPheValPheLeuValProGluThrIleGly 779
Db      912 GCAAGCTGTGCTGATGCTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 971
Qy      780 MetProLeuGluValIleThrGlnPheAlaValGlyAlaGlyAlaAlaAla 798
Db      972 ATGCCCTCGAGGTCATCACCGAGTTCTTTGGGGTGGGCGAAGCAGCGCAGGCC 1028

```

RESULT 12

US-10-051-909-13
Sequence 13, Application US/10051909
Publication No. US20020199217A1

GENERAL INFORMATION:

APPLICANT: Allen, Steve
APPLICANT: Helentjaris, Tim
APPLICANT: Hitz, Bill
APPLICANT: Kinney, Tony
APPLICANT: Tiney, Scott
TITLE OF INVENTION: Plant Sugar Transport Proteins
FILE REFERENCE: B01163 US CIP
CURRENT APPLICATION NUMBER: US/10/051,909
PRIOR FILING DATE: 2002-01-17
PRIOR APPLICATION NUMBER: 60/083,044
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Microsoft Office 97
SEQ ID NO 13
LENGTH: 1487
TYPE: DNA
ORGANISM: Trifolium aestivum
US-10-051-909-13

Alignment Scores:

Pred. No.: 3,36e-141 Length: 1487
Score: 1424.00 Matches: 281
Percent Similarity: 89.97% Conservative: 24
Best Local Similarity: 82.89% Mismatches: 32
Query Match: 34.94% Indels: 2

DB: 14 Gaps: 2
US-10-051-909-32 (1-800) x US-10-051-909-13 (1-1487)

```

Qy      462 GluIleGlyAspGlyValSerSerThrAspIleGlyGlyIleTrpGlnLeuAlaTrpIle 481
Db      12 GAGGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 71
Qy      482 TrpSerGluIleGlyGlyValGlyValGlyValGlyValGlyValGlyValGlyVal 501
Db      72 TGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 131
Qy      502 HisGlnIleGlyValProGlySerArgArgGlySerIleValSerLeuProGlyIleGly 521
Db      132 CACCAAGAGGGGGTGGCCGACTCAAGAGAGGGCTGTGTGTTCATCTTCGTTGGGGGT 191
Qy      522 AspValPheGlu---GlySerGluPheValHisAlaAlaLeuValSerGlnSerAla 540
Db      192 GATGCCACGCAAGGGGGCAGTGGGTTATATCATGCTGCTGTTGTAAGCCATCGGCT 251
Qy      541 LeuPheSerIleGlyLeuAlaGluProArgMet---SerAspAlaAlaMetValHisPro 559
Db      252 CTCTACTCTCAAGATCTTATGGAAGAGCGATGAGCGGCTCCAGCATGATTCATCCA 311
Qy      560 SerGluValAlaAlaIleGlySerArgTrpIleAspLeuPheGluProGlyValArgArg 579
Db      312 TTGAGGAGGAGCTCCCAAGGTTCAATCTGAAAGATCTTTTGAACCTGGTGTAGGCT 371
Qy      580 AlaLeuLeuValGlyValGlyIleGlnIleuGlnIleuPheAlaGlyIleAsnGlyVal 599
Db      372 GCATTGTCGTCGGTGTGGAATTCAGATGCTTCAAGCATGTTGCTGGAATAATGAGATT 431
Qy      600 LeuTYrTYrProGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 619
Db      432 CTCTACTACTCTCTCAATTCCTGAGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 491
Qy      620 GlyLeuSerSerAlaSerAlaSerIleuIleuIleuSerLeuThrThrLeuMetLeu 639
Db      492 GGCCTCAGTTCAGATCAGATCAGATCAGATCAGATCAGATCAGATCAGATCAGATC 551
Qy      640 ProCysIleGlyPheAlaMetLeuMetLeuMetLeuMetLeuMetLeuMetLeu 659
Db      552 CCAAGCATTTGGTGTGAGCATGAGCATTTATGATATCTGGAAGAAGTTTCTGTAAG 611
Qy      660 GlyThrIleProIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 679
Db      612 GGCACAATTTCCATCTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 671
Qy      680 LeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValIleValIle 699
Db      672 TTGAGTAGGAGGCCCCCAGCTGTGCTCTCCACAGTAGAGTGTATGCTACTCTGCTGC 731
Qy      700 PheValMetGlyPheGlyProIleProAsnIleuGlyValAlaGluIlePheProThrArg 719
Db      732 TTGTGCATGGGCTTTGGCCCATCCCAACATTCATGTGAGAGATTTCCCAACAGCA 791
Qy      720 ValArgGlyLeuGlyAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleVal 739
Db      792 GTCCGTGTGTCTGTGATGCTGATTTGGCCCTCCACATTCGTGATTTGAGATTTATGTT 851
Qy      740 ThrTYrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleIle 759
Db      852 ACCTACACCTGCGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 911
Qy      760 AlaValValCysLeuIleSerPheValPheValPheLeuValProGluThrIleGly 779
Db      912 GCAAGCTGTGCTGATGCTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 971
Qy      780 MetProLeuGluValIleThrGlnPheAlaValGlyAlaGlyAlaAlaAla 798
Db      972 ATGCCCTCGAGGTCATCACCGAGTTCTTTGGGGTGGGCGAAGCAGCGCAGGCC 1028

```

RESULT 13

```

US-10-051-902-15
; Sequence 15: Application US/10051902
; Publication No: US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tinney, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1009
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-051-902-15

Alignment Scores:
Pred. No.: 3.14e-90 Length: 1009
Score: 943.00 Matches: 179
Percent Similarity: 91.89% Conservative: 25
Best Local Similarity: 80.63% Mismatches: 18
Query Match: 23.14% Indels: 0
DB: 14 Gaps: 0

US-10-051-909-32 (1-800) x US-10-051-902-15 (1-1009)

QY 574 GIUPROGLYVALARGARGALALEULEUVALGLYALGLYLEGNILEUGINGINPHE 5939
Db 2 GAACCTGGAGTGAAGCATGCACACTGTTGTTGGCATGGATTACAGATCCGACAGAGTTT 61
QY 594 AIAAGLYLAAGGLYVALLEUPLYRTRTHPROGNILEUGINGINAGLYVALA 6133
Db 62 GGGGGATCAATGAAGATCCCTACTACACACCTCAATACTTACAGACAGAGGTGCGG 1212
QY 614 VALILEUSETLYSHEGLYLEUSERSERIALSERIALSERILEUULEUSERSERLEU 6333
Db 122 GTTCTTCAACAACATTGGACCTAAGCTCTTCTCAACATCTATTATTAGTCCCTTG 1818
QY 634 THRTRILEUUEULEUPROCYLIEGLYPHEALIMETLEUUEULEUETAPLEUSERGLY 6533
Db 182 ACAACCTTGCTGATGCTTCCAGCATTTGGCATGCCATGAGACTCATGATATGTACAGA 2414
QY 654 AGCARGPHELEULEUGLYTRHILEPROLILEUULELSEULEUVALILEUVAL 6733
Db 242 AAPAGGTTTCTTCTCTTCAACAATCCCTGCTTTATATACAGCTAGCTGTCTTGTT 3010
QY 674 VALSERASLEULELEASPLEUGLYTRHILEALHIALALEUUEUSERTRHVALSERVAL 6933
Db 302 TTAGTCGATCTTGTGATGTCGGAACCATGTCGACGCTCGCTTCAACGATCAGGCTC 3616
QY 694 ILEVALTYRPHCYCSPHEVALMETGLYPHEGLYPROLIEPROLEUVALCYEVAL 7133
Db 362 ATCGCTATTCTTCTGCTTTCGTCATGAGGGGTTGGCCCTATCCCAATATTCTTGCGCG 4212
QY 714 GLUILEPHEPOTHRARGVALAAGGLYLEUCYALALECYEVALAPHETRPHEPTR 7333
Db 422 GAGATTTTCCCAACCTCTGCTCCGTGCATGTCATGACCAATTCGCGGCTAACCTTCTGG 4818
QY 734 ILEGIAAPLILELEVALTHRYRSEULEUPROVALMETLEUENMIALILEGLYLEUVAL 7533
Db 482 ATCGGCGACATCAATCGACATACACTCTTCCCGTATAGTCATAGCCATTGGCTCTGCT 5411
QY 754 GLYVALPHESERIETRYALAVALCYSEULEUSERPHEVALPHEVALPHELEUYS 7733
Db 542 GGAGCTTTCGGCATATATAGCATCGTTTGTCATGACCTTTGATTCGCTACATGAG 6010
QY 774 VALPROGLUTHRILYSGLYMETPROLEUGLVALILETHRIGLUPHEALVALGLYVAL 7939

```

```

Db      602 GTCCCTGAGACAAAGGCGATGCCCTTGAGGTCATCACCGAATTTCTTCTGTGCGGGGCA 661
QY      794 Lyggin 795
      |||||
Db      662 AAGCAG 667

RESULT 14
US-10-051-909-15
; Sequence 15, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helenjavis, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Timney, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BBI163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051, 909
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083, 044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1009
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-051-909-15

Alignment Scores:
Pred. No.: 3.14e-90 Length: 1009
Score: 943.00 Matches: 179
Percent Similarity: 91.89% Conservative: 25
Best Local Similarity: 80.63% Mismatches: 18
Query Match: 23.14% Indels: 0
Db: 14 Gaps: 0

US-10-051-909-32 (1-800) x US-10-051-909-15 (1-1009)
QY      574 GIUPRQGIYVALARGARGALALEUENYALGIYVALGIYLEGINLEUGNGINPHE 593
      |||||
Db      2 GAACCTGAGAGAGACAGACGACACTGTCTTGGCATAGATTCACAGATCTCGACGACGTTT 61
QY      594 AIAgiYIAeangIYValIeunTYrThrProginIleleuGInaIaIaIa 613
      |||||
Db      62 GCGGGTATCATATGAGGTCCTCTACTACACACTCGATATCTTGAGCAAGCAGGTGCGGG 121
QY      614 ValIleuSerLYsPhgGIYleuSerSerAlaSerIleleuIleSerSerLeu 633
      |||||
Db      122 GTTCTTCTATCAAAACATTCGACTAAGCTCTTCTCAGCATCTATTTCTTATGTGCTTG 181
QY      634 ThrThrIleuLeuMetLeuProCysIleGIYPhaIaMetLeuLeuMetAspLeuSerGIY 653
      |||||
Db      182 ACAACCTTGCGATGATCTTCCAGCATGGCATGCCCATAGACTCATGGATATGTACAG 241
QY      654 ArgArgPheLeuLeuLeuGIYThrIleProIleuIleAlaSerLeuValIleuVal 673
      |||||
Db      242 AGAAGGTTCTTCTTCTCTTCAACATCTCCTGTGATATGATAGCGGTACTGTGTTGGTT 301
QY      674 ValSerAsnLeuIleAspLeuGIYThrIleuAlaHISAlaLeuLeuSerThrValSerVal 693
      |||||
Db      302 TTAGTGATGTGTTCTGATGTGGAAACATATGGGCAAGCGCGCTCTCAACGATCAACGCTG 361
QY      694 IleValTYrrPheCysCysPheValMetGIYPhgGIYProIleProAsnIleuCysAla 713
      |||||
Db      362 ATCGCTATTTCTGTCTTCTGTCTCATGGGGTTTGCGCTATCCCAAAATATTTCTGTGCGG 421
QY      714 GluIlePheProThrArgValArgIYleuCysIleAlaIleCysAlaPheThrPheP 733
      |||||
Db      422 GAGATTTTCCCACTCTGTGTCCGTGGCATCTGCATATGCCATCTGCGCGCTTAACTTCTGG 481

```

Qy 734 IleglyaspriilelevalthrtyrserleuProValMetLeuAsnAlaIleglyLeuAla 753
Db 482 ATCCGCGCATCATCGTACATACCTCTCCCGGATGCTCAATGCGATCGTCTGCT 541
Qy 754 GlyValPheSerIleTyrAlaValAlaValCysLeuIleSerPheValPheValPheLeuLys 773
Db 542 GGAGCTCTGGCGCAATATGCGATCGTTGTGTACTAGCCCTTGTATGTCGTACATGAAG 601
Qy 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
Db 602 GTCCCTGAGCAAAAGGCGATGCCCTGGAGGCTCATACCGAGCTTCTCTGTGCGGCA 661
Qy 794 LysGln 795
Db 662 AAGCAG 667

RESULT 15
US-10-369-493-47011
Sequence 47011, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 47011
LENGTH: 1374
TYPE: DNA
ORGANISM: Bacillus subtilis
US-10-369-493-47011

Alignment Scores:
Pred. No.: 3,97e-52 Length: 1374
Score: 587.50 Matches: 168
Percent Similarity: 36.64% Conservative: 102
Best Local Similarity: 22.80% Mismatches: 162
Query Match: 14.42% Indels: 305
DB: 12 Gaps: 12

US-10-051-909-32 (1-800) x US-10-369-493-47011 (1-1374)

Qy 55 LysMetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGly 74
Db 4 AAAAAGCAGTCAMATATATGCTTATATTTTCGAGGCTCTGGAGCGCGCTTATATATGCG 63
Qy 75 TrpAspAsnAlaThrIleAlaAlaValLeuTyrIleLysGluPheGlnLeu 94
Db 64 TATGATACCGAGATATTCGAGCTATTTATTTATGAAAAAGAGTAGGCTTA--- 120
Qy 95 AsnGluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleVal 114
Db 121 ---AACGCGTTTACAGAGGCTCTGTGTCAGCTCTTGCTGGGCGCATATTGGAGC 177
Qy 115 ThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgPrometLeuIleLeuSer 134
Db 178 TCAGAGGCGCGCGAGCTGACTGACCTGTTGGAAGAAAAAACAATTATGGACGCC 237
Qy 135 SerIleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeu 154
Db 238 GCGCTGCTGTTTGTATAGCGGCTCTGTGTGCGACATGCGCCCAATACAGAGTCAAG 297
Qy 155 LeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeu 174
Db 298 GTGCTGTTTCGATCATTTTGGAGCTTGCAATCGAACAATCGACATCATCCCTT 357

Qy 175 TyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPhe 194
Db 358 TATTATCTAGACTGCGGCCAAACATAAAGCGGCGCTGTATCATCATCATCAGCTG 417
Qy 195 SerGlySerGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerPro 214
Db 418 ATGATCAAGCGGCGATCTCTTCTTAC-----ATTGCATATCATATTGCGCAT 471
Qy 215 SerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerIlePhePhePheGly 234
Db 472 GCCAAGCGTGGCGCTGATGCTTGTGATGCTGCTGTGCGTATGCTCTTCTTCTT--- 528
Qy 235 LeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAla 254
Db 529 ATTGCAATTTGTTTATGCGCGAGAGCGCGCTGCTGTACAGATGCGGAAGAAAGC 588
Qy 255 GluAlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSer 274
Db 589 AAACGAAAGAAATTTGAAATAATTCGCTGCGACAAAGATATT----- 633
Qy 275 LeuLeuLeuGluGlyLeuGluValGlyLysPheThrSerIleGluGluTyrIleGly 294
Db 633 ----- 633
Qy 295 ProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeu 314
Db 633 ----- 633
Qy 315 TyrGlyProGluGluGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeu 334
Db 633 ----- 633
Qy 335 GlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeu 354
Db 633 ----- 633
Qy 355 MetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGly 374
Db 633 ----- 633
Qy 375 SerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHis 394
Db 633 ----- 633
Qy 395 AlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSer 414
Db 634 -----GATCAGGA----- 642
Qy 415 AspGlyAlaGlyLysPheTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla 434
Db 642 ----- 642
Qy 435 ThrGlyAlaGluGlyLysAspIleValHisIleGlyHisArgGlySerAlaLeuSerMet 454
Db 643 -----ATCAT----- 648
Qy 455 ArgArgIleThrLeuLeuGlyGluGlyLysAspGlyValSerSerThrAspIleGlyGly 474
Db 649 -----GATATA----- 654
Qy 475 GlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGluGluAsnGlyArgLysGluGly 494
Db 655 -----AAAGAAAGCGAAAGAGCATGAAAGCGCGCTGAAAGAG--- 693
Qy 495 GlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIle 514
Db 693 ----- 693
Qy 515 ValSerLeuProGlyGlyLysAspValPheGluGlySerGluPheValHisAlaAla 534
Db 693 ----- 693

```
Oy 535 LeuValSerGlnSerAlaLeuPheSerIleuGlyLeuAlaGluProArgMetSerAspAla 554
Db 693 ----- 693
Oy 555 AlaMetValHisProSerGlnValAlaAlaGlySerArgTrpIleAspLeuPheGlu 574
Db 694 -----CTGTTCGAT 702
Oy 575 ProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAla 594
Db 703 CCAATGGGTGCGCCCGACGCTTATTCAGAGTTGGAGCTCGCTTTTTCAGCAATTATTC 762
Oy 595 GlyIleAsnGlyValLeuTyrtTrpProGlnIleLeuGlnAlaGlyValAlaVal 614
Db 763 GGAACGAATACATCATCTACTATGCGCAAGACCTTACAACGTCGGA----- 813
Oy 615 IleLeuSerLysPheGlyLeuSerSerAlaSerIleLeu-----Ile 630
Db 814 -----TTCGGA-----AACTCCGCTTCGATTTTAGCGCACGCTCGGAATC 852
Oy 631 SerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAsp 650
Db 853 GGCACAGTCATATGTTCTCATGACATTA-----GTAGCGATTAATAATCATCGAC 900
Oy 651 LeuSerGlyArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuVal 670
Db 901 AAGATTGGAGAGAACCCGTTACTGCTATTCGGGAATGCGGCAATGATGATCAGCTTGATC 960
Oy 671 IleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThr 690
Db 961 GTTTCGCTTAAATCTCTTTTCAATACACTCCGCTCCCTCATGACGACCGTTC 1020
Oy 691 ValSerValIleValIlePheCysCysPheValMetGlyPheGlyProIleProAsnIle 710
Db 1021 ATTGTATTAGGCGTGTATTATCGTGTCTTTCGCGTCACTGCGGACCGGTGTGGGTG 1080
Oy 711 LeuCysAlaGluIlePheProThrArgValAlaArgIleuGlyLeuAlaIleCysAlaPhe 730
Db 1081 ATGCTTCCTGAATTTGCCGCTTCAGCGTCAAGAGAAATCGGACCGGTGTTGCACCTTA 1140
Oy 731 ThrPheTrpIleGlyAspIleIleValThrTyrtSerLeuProValMetLeuAsnAlaIle 750
Db 1141 ATGTTCGACGTTGGACACTGATTTGTTCAATACCTATCCAAATTAATATGAGACCGATC 1200
Oy 751 GlyLeuAlaGlyValPheSerIleTyrtAlaValValCysLeuIleSerPheValPheVal 770
Db 1201 GGAATCACTATTATTCTCTGATTATAGCCGATCGGATCATGCGCTTATTATTGTC 1260
Oy 771 PheLeuLysValProGluThrLysGlyMetProLeuGluValIleThrGlu 787
Db 1261 CGATTTAAGTGACAGACAGAAAGGAAAGCCTTGAAGAAATTGACGACG 1311
```

Search completed: January 6, 2004, 00:41:21
Job time : 748 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 5, 2004, 19:06:39 ; Search time 122 Seconds

(without alignments)
2894.316 Million cell updates/sec

Title: US-10-051-909-32

Perfect score: 4075
Sequence: 1 IRSGSWLAVTFPPDLDR.....PLEVITEFPAVGAQAAAKA 800

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

```
MODEL=frame_p2n.model -DEV-xlh
-O=/cgn2_1/USPTO_spool/US10051909/runat_05012004_094743_22329/app_query.fasta_1.967
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdt
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pro -NOR=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10051909@cgn2_1.1.56@runat_05012004_094743_22329 -NCP=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database : Issued_Patents_NA:*

```
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/6C.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/6D.COMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3517	86.3	2824	4	US-09-291-922-1
2	2674	65.6	2601	4	US-09-291-922-7
3	1670.5	41.0	1692	4	US-09-291-922-9
4	1424	34.9	1487	4	US-09-291-922-13
5	943	23.1	1009	4	US-09-291-922-15
6	562	13.8	870	4	US-09-291-922-5
7	496	12.2	1853	4	US-09-291-922-23
8	494.5	12.1	2017	4	US-09-291-922-21
9	491.5	12.1	1914	4	US-09-291-922-19
10	489	12.0	1872	4	US-09-291-922-27
11	463.5	11.4	2089	4	US-09-291-922-25
12	412	10.1	4411529	3	US-09-103-840A-1

13	368	9.0	4403765	3	US-09-103-840A-2	Sequence 2, Appli
14	355	8.7	2856	4	US-09-643-597-135	Sequence 135, App
15	355	8.7	2856	4	US-09-480-884A-135	Sequence 135, App
16	355	8.7	2856	4	US-09-542-615A-135	Sequence 135, App
17	355	8.7	2856	4	US-09-606-421B-135	Sequence 135, App
18	354.5	8.7	3000	2	US-08-928-692-9	Sequence 9, Appli
19	354.5	8.7	3000	4	US-09-339-972-9	Sequence 9, Appli
20	332.5	8.2	2592	4	US-09-591-025-8	Sequence 8, Appli
21	275.5	6.8	972	4	US-09-114-001C-1798	Sequence 1798, Ap
22	266.5	6.5	443	4	US-09-291-922-3	Sequence 3, Appli
23	265.5	6.5	1428	4	US-09-328-352-1580	Sequence 1580, Ap
24	256	6.3	918	4	US-09-107-532A-561	Sequence 561, App
25	247	6.1	2343	2	US-09-031-392-1	Sequence 1, Appli
26	247	6.1	2343	3	US-09-299-549-1	Sequence 1, Appli
27	247	6.1	2343	4	US-09-610-417-1	Sequence 1, Appli
28	235	5.8	155	4	US-09-313-234A-4597	Sequence 4597, Ap
29	234.5	5.8	510	4	US-09-291-922-11	Sequence 11, Appli
30	227.5	5.6	1467	4	US-09-328-352-2465	Sequence 2465, Ap
31	225.5	5.5	1521	4	US-09-252-991A-989	Sequence 989, App
32	218.5	5.4	1002	4	US-09-252-991A-1099	Sequence 1099, Ap
33	217.5	5.3	1371	4	US-09-328-352-1320	Sequence 1320, Ap
34	207.5	5.1	1230	4	US-09-252-991A-947	Sequence 947, App
35	205.5	5.0	603	4	US-09-252-991A-1063	Sequence 1063, Ap
36	205	5.0	2842	4	US-09-620-312D-574	Sequence 574, App
37	198.5	4.9	1344	4	US-09-252-991A-3992	Sequence 3992, Ap
38	198.5	4.9	1479	4	US-09-252-991A-3960	Sequence 3960, Ap
39	196.5	4.8	2102	2	US-08-647-397-1	Sequence 1, Appli
40	192	4.7	1896	3	US-08-501-572-6	Sequence 6, Appli
41	192	4.7	1896	3	US-09-040-444-6	Sequence 6, Appli
42	190.5	4.7	2871	4	US-09-252-991A-11431	Sequence 11431, A
43	190	4.7	1320	4	US-09-252-991A-10535	Sequence 10535, A
44	190	4.7	1473	4	US-09-252-991A-10639	Sequence 10639, A
45	187.5	4.6	1251	4	US-09-328-352-2549	Sequence 2549, Ap

ALIGNMENTS

```
RESULT 1
US-09-291-922-1
; Sequence 1, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 2824
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (29)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (622)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (636)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (638)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (669)
```

```

FEATURE:
NAME/KEY: unsure
LOCATION: (771)
FEATURE:
NAME/KEY: unsure
LOCATION: (822)
FEATURE:
NAME/KEY: unsure
LOCATION: (856)
FEATURE:
NAME/KEY: unsure
LOCATION: (889)
FEATURE:
NAME/KEY: unsure
LOCATION: (896)
FEATURE:
NAME/KEY: unsure
LOCATION: (944)
US-09-291-922-1

```

```

Alignment Scores:
Pred. No.: 0
Score: 3517.00
Percent Similarity: 95.18%
Best Local Similarity: 91.57%
Query Match: 86.31%
DB: 4
Gaps: 2
Length: 2824
Matches: 684
Conservative: 27
Mismatch: 34
Indels: 2

```

US-10-051-909-32 (1-800) x US-09-291-922-1 (1-2824)

```

QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr 75
DB 238 ATGGGGGGGGCGCTGATGGTCCGCATCGGGGCTCTATCGCACTGCTGCAGGGCTGG 297
QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleValysGluPheGlnLeuGlnAsn 95
DB 238 GACAAATGCAACATTCGTGAGCCGCTCTGATCAATAGAGAAATTCACCTGCAGAGC 357
QY 96 GluProthValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
DB 358 GAGCCTCTGATGAGGCTCATCGCTGATGTTCTCATTTGGGCAACAGTCATCACA 417
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
DB 418 ACATCTCCGGGGCCAAAGGCTGACTGCTGTGGTAGAGGCCCATGTGGTCCCTCGCT 477
QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu 155
DB 478 GTCTCTACTTCGTGAGTGGCTGTGGTGTGGTGGCCCAATGTGTACATCTTGTCTC 537
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
DB 538 CTGGAGAGGCTCATGATGAGTTCGGTATCGGTTGGCCGTCACCTGTCTCTAC 597
QY 176 IleSerGluIleAlaProSerGluIleArg--GlyLeuLeuAsnThrLeuProGlnPhe 194
DB 598 ATCTCGAAGACTGCACCCGACAGAAATTTTGGGGCTGTTNAGAACAGTTCGCGAGTTC 657
QY 195 SerGly--SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213
DB 658 ATTTGGGGTCAAGAGGAGGATGTTCTCTCTCTGATGATGTTGGATGTCCTTCATG 717
QY 214 ProSerProAspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhePhe 233
DB 718 CCCAAACCTGATGAGGCTCATGCTTGGATGTTCTGATCCCGTCACTTATNTACTTT 777
QY 234 GlyLeuThrIlePheTyrLeuProGlnSerProArgTyrLeuValSerIleGlyArgMet 253
DB 778 GACATGACTGCTTCTTACTTCTGCTGATACCAAGGTGCTGTNAGCAAGAGAGATG 837
QY 254 AlaGluAlaIleValLeuGlnIleLeuValArgIleValAspAspValSerIleGlyLeu 273
DB 838 GCGGAGGCGAAGAGATGNTGCAAGAGCTGCGGGGAGAGAGAAATGTCTCAGGAGANG 897

```

```

QY 274 SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIle 293
DB 898 GCTCTTCTAGTGAAGGTTTGGGGTCCGTAAGATACCTATTTNAGATACATCTT 957
QY 294 GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspGlyGlnIleThr 313
DB 958 GGAAGTGCACGAGGAGCCGATGATCTTGAATCAAGGTGATTAAGAAACAATCA 1017
QY 314 LeuTyrGlyProGluGluGlnSerThrIleAlaArgProSerLeuGlyProIleMet 333
DB 1018 CTTATGGGCTGGAAGAGGCGAGTATGCTGACCTTCAAGGAGCCATCATG 1077
QY 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
DB 1078 CTGGAGATGTCCTTCTTCTGCACTTCGTATGGAGACATGTCAGAACAGATGACC 1137
QY 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
DB 1138 CTTATGATCCGATGTGACACTTTTGGTGTGTCATGAGAAATATGCTCAAGCTGGA 1197
QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
DB 1198 GGAAGTATGAGAGACATTTGTTCCAAACTTGGAGTATGTTCAAGTGCACAGATCAG 1257
QY 394 HisAlaIysAsnGluGlnIleThrAspGluGluAsnLeuHisArgAspAspGluTyrAla 413
DB 1258 CATGCCAAATATGACAGTGGAGATGAAAGATCTTCAATAGAGATGAGAGAGATACCA 1317
QY 414 SerAspGlyAlaGlyGlyAspTyrGlyAspAsnLeuHisSerProLeuLeuSerArgGln 433
DB 1318 TCTGATGTCAGAGAGAGTGAATATAGAGACATCTCCATACCATTCCTTCCAGGAG 1377
QY 434 AlaThrGlyAlaGluGlyValAspIleValHisIleGlyHisArgGlySerAlaLeuSer 453
DB 1378 GCAACAGGTGCGGAGGAGAAAGACATTCGTGACCATGTCACCTGGAAAGCTTTGAGC 1437
QY 454 MetArgArgGlnIleThrLeuGluGlyGlyValAspGlyValSerSerThrAspIleGly 473
DB 1438 ATGAGAGGCAAGGCTCTTGGAGAGGTGAGATGTTGAGAGCAGCATGATATCGT 1497
QY 474 GlyGlyTyrGlnLeuAlaIleTyrLeuTyrSerGluIleGlyGlyAsnGlyArgGlyGlu 493
DB 1498 GGGGATGAGCAGCTTGTGTAATGTCAGAGAAAGGTGAGAAATGTAAGAAAGAA 1557
QY 494 GlyGlyPheIleValGlyValTyrLeuHisGlnGluGlyValProGlySerArgArgIleSer 513
DB 1558 GGTGTTTCAAAAGATCTTACCTGCACCAAGAGGAGTTCCTGCTCAAGAGGGGCTCA 1617
QY 514 IleValSerLeuProGlyGlyValAspValPheGluGlySerGluPheValHisAlaAla 533
DB 1618 ATGTTTCACTTCCCGGTGAGGAGATGTTCTTAAGGATGATGATTTGTATACGTGCT 1677
QY 534 AlaLeuValSerGlnSerAlaLeuPheSerIleGlyLeuAlaGluProArgMetSerAsp 553
DB 1678 GCTTTAGTAATCACTGACGACTTTTCTCAAGAGGTCTGTGTAACACGATGTCAGAT 1737
QY 554 AlaAlaMetValHisProSerGluValAlaAlaIleGlySerArgTyrPheAspLeuPhe 573
DB 1738 GCTGCATGCTTCAACCATCTGAGATGATGCTCCAAAGGTTCACCTTGGAAAGATTTGTT 1797
QY 574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPhe 593
DB 1798 GAACCTGAGAGAGGCGCTTGTGTCGTTGTGAAATTCAGATCTTCAACAGATT 1857
QY 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGlnAlaGlyValAla 613
DB 1858 GCTGGAATTAACGAGTGTCTTACTATACCCACAATTTCTTGAAGAGAGTGTGCA 1917
QY 614 ValIleLeuSerIlePheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
DB 1918 GTTATCTTTCCAATTTGCTCAGCTCGGATCAGATCATCTGATCATGATTCCTCTC 1977

```



```

Db      1297 ACAACAGCCTGAA---AAAGCTTACTCCTCCTCCTCCATCGACGATTCCTGGC 1353
Qy      453 SerMetArgArgGlnThrLeuLeuGlyGly---GlyAspGlyValSerSerThrAsp 471
Db      1354 AGCATGAGGCGTCACAGTGTCTCATGCAAGGGTCAGGAGCAAGGTGTGTACAGCT 1413
Qy      472 ILeGlyGlyGlyTyrGlnLeuAlaTyrPlySTPserGlnuysGlnGlyValuansGlyArg 491
Db      1414 ATTGGTGGGCTGGCACTGGCATGGAATGAGCTGATAAA---GGTAGAGTGAANA 1470
Qy      492 IysGlnGlyGlyPheLeuArgValTyrLeuHISglnGlnGlyValProGlySerArgArg 511
Db      1471 CAACAGAGGAGGCTTTAAAGATTTATTTACATGAGAGGAGGTTCTGCATCTGTGCT 1530
Qy      512 GlySerIleValSerLeuProGlyGlyGlyAspValPheGlnGlySerGlnuysValHis 531
Db      1531 GGATCCATTGTATTCATTCCTCCGTGAAAGC-----GAATTTGTCCAG 1572
Qy      532 AlaAlaAlaLeuValSerGlnSerAlaLeuPheSerGlyLeuAlaGlnProArgMet 551
Db      1573 GCTGCTGCTTGTGTAAAGCCACCCGCTTTTACTCTAAGAGACTTATTTGATGACACCA 1632
Qy      552 SerAspAlaAlaMetValHisProSerGlnuysAlaAlaAlaGlySerArgTyrPlyAsp 571
Db      1633 GTTGGGCGCTGCAATGGTTCACCCATCTGAGACAGCTTCAAGGGGCAAGTTGGAAGCT 1692
Qy      572 LeuPheGlnProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln 591
Db      1693 CTCTTGAACCAAGGGGTTAAGCATGTATGTGTGGAGTTGGAAATCAAAATACCTTCG 1752
Qy      592 GlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGlnGlnAlaGly 611
Db      1753 CAGTTTTCAGGATTAATGGGTTCTATTTTACACCTCAATCTTGAAGAGCGGCT 1812
Qy      612 ValAlaValIleLeuSerIlePheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer 631
Db      1813 GTTGAAGTCTCTTTCAGATATAGGCATGTGGCTGAGTGGCATCATCTTTATTCAGT 1872
Qy      632 SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu 651
Db      1873 GCTTTCACAACTTCTGTATGTCTTCCCTGTATAGGGGTAGCCATGAGTCAATGATGT 1932
Qy      652 SerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle 671
Db      1933 TCAGGCAGAGGACAGTGTCTTACTTACTACATCCCGTCTGATTTGTGTCACCTCATATT 1992
Qy      672 LeuValIleSerAsnLeuIleAspLeuGlyThrIleuAlaHisAlaLeuLeuSerThrVal 691
Db      1993 TTGGTCATTGGAAGCTGGTAAATTTTGGCAATGTGCCCATGCAGCAATCTCAACAGTA 2052
Qy      692 SerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu 711
Db      2053 TGGCTGTGGTATTTATTTCTGCTGTGTGTGTATGAGGTTATGACCAATTCCAAACTCCT 2112
Qy      712 CysAlaGlnIlePheProThrArgValArgGlyLeuCysAlaIleCysAlaIlePheThr 731
Db      2113 TGCTAGAGATTTTCCCACTAGGGGTGGCTGTGATGCTGATTCGATCTGTGCTATTAAG 2172
Qy      732 PheTyrIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGly 751
Db      2173 TTCTGATTTGAGACATCATCATACATCTGCTGCTGCTGATGCTGGCTCTTTAAGA 2232
Qy      752 LeuAlaGlyValPheSerIleTyrAlaValCysLeuIleSerPheValAlaPhe 771
Db      2233 CTGGTGTGGTATTTCCATTTACGACAGTGTGTGTGTATCTCGTGGATATTTGTGTT 2292
Qy      772 LeuIysValProGlnThrIysGlyMetProLeuGlnuysIleThrGlnPhePheAlaVal 791
Db      2293 TTGAAGTTTCCAGAAACAAAGGCAAGCCCTTGAAGTATCTGTGAATTTCTTTCTGTT 2352
Qy      792 GlyAlaIysGlnAlaAlaVal 798

```

```

Db      2353 GGAGCAAGACAGCTGCTTCT 2373
RESULT 3
US-09-291-922-9
; Sequence 9, Application US/09291922
; Patent No. 6383776
GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Timney, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 1692
; TYPE: DNA
; ORGANISM: Glycine max
US-09-291-922-9

Alignment Scores:
Pred. No.: 1,53e-161 Length: 1692
Score: 1670.50 Matches: 333
Percent Similarity: 79.75% Conservative: 57
Best Local Similarity: 68.10% Mismatches: 84
Query Match: 40.99% Indels: 15
DB: 4 Gaps: 8

US-10-051-909-32 (1-800) x US-09-291-922-9 (1-1692)
Qy      308 AsPlyGlnGlnIleThrLeuTyrGlyProGlnGlnGlyGlnSerTyrIleAlaArgPro 327
Db      21 GAATAAGTCAATTAATTAGTTATGAGACCAAGAGCCAGTGGTGTGTAAGACT 80
Qy      328 SerIysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMet 347
Db      81 GTTGTGAGCAACCAATTTCTGTGGC-----CTTGTATCTAGCAAGAAAGCAATG 128
Qy      348 ValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGly 367
Db      129 GCAATTCACAC---AGCTTGTGAGACCTCTTATGACCTCTTTGGTATGATACATAG 185
Qy      368 AsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMet 387
Db      186 AAGTCCCAAGAAACAGG-----AGCACCTTTTTCACACTTGGGAGATAG 233
Qy      388 PheSerValThrAspGlnHisAlaIleAsnGlnGlnIleThrAspGlnGlnuysAsnLeuHis 407
Db      234 TTTCAGTGTGGGGAATACACCAAGAAATGAATGGATGGAGAAAGCTTGAAGCCGA 293
Qy      408 AspAspGlnGlnuysTyrAlaSerAspGlyValaGlyIysAspTyrGlnuysAsnLeuHisSer 427
Db      294 GAGGTATGATTTATGTTCTGAT-----CTGTGATTTGTGATGACAAATTTGCAAGT 347
Qy      428 ProLeuLeuSerArgGlnAlaThrGlyAlaGlnGlyIysAspIleValHisGlyHis 447
Db      348 CCAATTGATCTACGTCACCAAGACAGGTGTGAT---AAGACATATCCCTCCATGCCAT 404
Qy      448 ArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu---GlyGlnGlyIysAspGly 466
Db      405 AGTAACCTTGA---AGCATAGGCAAGTATCTTTTACATGGAATTCAGAGAACCC 461
Qy      467 ValSerSerThrAspIleGlyGlyIleThrGlnLeuAlaTyrPlySTPserGlnuysGln 486
Db      462 ACTGATGATCTGAGATTTGGTGGTGGAGCTGACATGGAATGATGCTGAAGAGAG 521
Qy      487 GlyIysGlnGlyArgIysGlnGlyIlePheIysArgValTyrLeuHisGlnGlnGlyVal 506

```

```

Db 522 GGGCCAGATGAAAGAGAGAGTGGCTTCAAGATAATATTTTACACCAAGATGGTGT 581
Qy 507 ProglySerArgArgGlySerIleValSerLeuProGlyGlyValAspValPheGluGly 526
Db 582 TCTGATCTTAAGACGTGGGTCTGTGGTTTCACTCCCT---GGGGGAGATTACCACTGAC 638
Qy 527 SerGluPheValHisAlaAlaValSerGlnSerAlaIlePheSerLysGlyLeu 546
Db 639 AGTAGAGTTGTAACAGCTGCTGCTGGTGAAGTCAAGCCCTTTATTAATAGAACCTT 698
Qy 547 AlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaGly 566
Db 699 ATGGCTCAACGCCAGTTGACCAAGCTATGATTCATCCCTCGAAACAATTGCCAAGAGG 758
Qy 567 SerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGly 586
Db 759 CCAAGTTGAGTGAATCTTTTGAACCTGGGATGAAGCATCATATATGTGGGGGTGGGA 818
Qy 587 IleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTrpTrpProGlnIle 606
Db 819 ATGCAATTTCTTCAAGATTCCTGTGTAATATGGGGTCTTACTATACGCTCAATTT 878
Qy 607 LeuGlnGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAla 626
Db 879 CTTGAGCAGGCGAGGTGGTGTATCTTCTTCAAGCCTAGAGCCTTGTTCTTCTTCA 938
Qy 627 SerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProGlyIleGlyPheAlaMet 646
Db 939 TCTCTTCTTATTAATGAGCGGTGACCAACTGTGTGAAGCTTCTTGTATAGCCATTGCCAG 998
Qy 647 LeuLeuMetAspLeuSerGlyValArgArgPheLeuLeuLeuGlyThrIleProIleLeu 666
Db 999 AGGCTCATGAGATTTTCAAGCGAAGAGACTTGGCTGCTCAAGACATCCCTCTTAATA 1058
Qy 667 AlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAla 686
Db 1059 GCAGCTCTTCTCATATTAAGTCTGCGGAAGCTTGTGGATTTGGATTCACCTGCAATGCA 1118
Qy 687 LeuLeuSerThrValSerValIleValIleValIleValIleValIleValIleValIle 706
Db 1119 TCACCTCAACCAATTAAGTGTATATGCTATATCTATCTTCTTGTCAATGGGATTTGACCA 1178
Qy 707 IleProAsnIleLeuCyAlaGluIlePheProThrArgValArgGlyLeuCyAlaIleAla 726
Db 1179 ATTCTATATATATCTTGGACAGAGATCTTCCCACTCAAGTCTGGTGTCTGCACTTGT 1238
Qy 727 IleCyAlaPheThrPheTrpIleGlyAspIleIleValIleThrTrpSerLeuProValMet 746
Db 1239 ATTGTGCTTACCTTGTGGATCTGTGATATCATTTGTCACCTACCACTCCAGTTAAG 1298
Qy 747 LeuAsnAlaIleGlyLeuAlaGlyValPheSerIleIleValAlaValCysLeuIleSer 766
Db 1299 CTCAAATTCGTAAGGCTCGCTGCTGTATTGATTAACTGCTGTGCTTCAAGCA 1358
Qy 767 PheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThr 786
Db 1359 TGGGCTTTGTCTTTTGAAGTTCAGAAACCAAGGCGATGCCACTGGAATGATCATT 1418
Qy 787 GluPhePheAlaValAlaGlyAlaLysGln 795
Db 1419 GAGTCTTCTCTGTGCGAGCAAAACAG 1445

```

```

; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1487
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-291-922-13

Alignment Scores:
Pred. No.: 2,61e-136 Length: 1487
Score: 1424.00 Matches: 281
Percent Similarity: 89.97% Conservative: 24
Best Local Similarity: 82.89% Mismatches: 32
Query Match: 34.94% Indels: 2
DB: 4 Gaps: 2

US-10-051-909-32 (1-800) x US-09-291-922-13 (1-1487)

Qy 462 GluGlyGlyAspGlyValSerSerThrAspIleGlyGlyIleValIleValIleValIleVal 481
Db 12 GAGGCTGGGAGGAGAGTCAAGCAGCATGCTATTTGGTGGGGGTGGCACTGCAATGGAA 71
Qy 482 TrpSerGlyuSerGluGlyValAsnGlyValArgGlyGlyIleValIleValIleValIleVal 501
Db 72 TGGTGGAGCACAAGGAGGAGATGGCAAGAGAGAGGCTTCAAAAGATCTACTTTG 131
Qy 502 HisGlnGlnGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGly 521
Db 132 CACCAAGAGGGGGTGGCCGAGTCAAGAGAGGGCTGTGTTTCACTTCTGGTGGGGGT 191
Qy 522 AspValPheGlu---GlySerGluPheValHisAlaAlaValSerGlnSerAla 540
Db 192 GATGCAAGCAAGGGGGAGTGGGTTTATACATGCTGCTGTTGGTAAGCCACTCGGCT 251
Qy 541 LeuPheSerLysGlyLeuAlaGluProArgMet---SerAspAlaAlaMetValHisPro 559
Db 252 CTTAATCTCAAGATCTTATGAAAGAGGTATGGGGCGGCTCCAGCATATATTCATCA 311
Qy 560 SerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArg 579
Db 312 TTGAGAGCAGCTCCCAAGGTTCAATCTGAAAGATCTGTTGAACCTGTGTAGGGCT 371
Qy 580 AlaLeuLeuValGlyValGlyIleGlnIleLeuGlnIlePheAlaGlyIleAsnGlyVal 599
Db 372 GCATTGTCGTGGGTGGATTCAGATGCTTCAGCAGTTTGTGGAATTAATGAGATT 431
Qy 600 LeuTrpTrpThrProGlnIleLeuGlnAlaGlyValAlaValIleLeuSerLysPhe 619
Db 432 CTCACTATATCTTCAAAATCTGGAGCAACCTGTGGTGTCTTCTTCAATCTT 491
Qy 620 GlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeu 639
Db 492 GGCCTCAAGTTCAGATTCAGATTCATCTTGATCACTTCTTCAACCACTTACTATGCTC 551
Qy 640 ProCyIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeu 659
Db 552 CCAAGCATTTGGTGTGACCATGAGACTTATGATATATCGAAGAAAGGTTTCTGTACTG 611
Qy 660 GlyThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAsp 679
Db 612 GGCACAAATTCATCTTATAGCATCCCTAATTTGTTGGGTGTGTGCAATGTAATCAAC 671
Qy 680 LeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValIleValIleValIle 699
Db 672 TTGAGTACGGTGGCCCAAGCTGTGTCTTCCACAGTACGTCATTTGTTCTTCTGCTGC 731
Qy 700 PheValMetGlyPheGlyProIleProAsnIleLeuCyAlaGluIlePheProThrArg 719
Db 732 TTTGTCAATGGCTTTGGCCGATCCCAACATTTCTATGTGACAGATTTTCCCAACAGA 791

```

```

RESULT 4
US-09-291-922-13
; Sequence 13, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Kinney, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163

```

QY 720 ValArgGlyLeuCyseIleIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleVal 739
Db 732 GTCCGTGGTGTGCTGATCCCTATTTCGCCCTCACATTGCGATTGTGACATTATTGTT 851
QY 740 ThrTyrSerLeuProValMetLeuAlaIleGlyLeuAlaGlyValPheSerIleTyr 759
Db 852 ACCCTACAGCTCCTGTGATGCTGAATGCTAATGGTCTAGCGGGTGTCTTTGTGTAATAT 911
QY 760 AlaValAlaCysLeuIleSerPheValPheValPheLeuIleValProGluThrIleGly 779
Db 912 GGAAGTGTGTGCTGATGCTTGTGCTGCTACCTAAAGCTCCAGACGCAAGAGGCC 971
QY 780 MetProLeuGluValIleThrGluPhePheAlaValAlaGlyAlaGlyAlaIleVal 798
Db 972 ATGCCCTCGAGCTATCATCCAGATTCTTTCGGCTTGGGGCCAGACGCGAGGCC 1028

RESULT 5
US-09-291-922-15
; Sequence 15, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1009
; TYPE: DNA
; ORGANISM: Trifolium aestivum
US-09-291-922-15

Alignment Scores:
Pred. No.: 3,566-87 Length: 1009
Score: 943.00 Matches: 179
Percent Similarity: 91.89% Conservative: 25
Best Local Similarity: 80.63% Mismatches: 18
Query Match: 23.14% Indels: 0
Gaps: 0
DB: 4

US-10-051-909-32 (1-800) x US-09-291-922-15 (1-1009)

QY 574 GluProGlyValArgAlaLeuValGlyValGlyIleGlnIleLeuGlnIlePhe 593
Db 2 GAACCTGGAGTGAAGATGACACTGTTGCTGGCATRAGATTCAGATCTGACAGATT 61
QY 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGlnIleValAla 613
Db 62 GCGGATACATAGAGTCTCTACTACACACTCAGATCTGAGACAGAGTGTCCGG 121
QY 614 ValIleLeuSerIlePheGlyLeuSerSerIleAlaSerIleLeuIleSerIleu 633
Db 122 GTTCTTCAATCAACATTGGACTTCTCTCTCACTCACTTCTTATTTAGTCCCTTG 181
QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
Db 182 ACAACCTTGCTATGCTTCTCCAGCATGGCATGGCATGGCATGGCATGGCATGGCATGG 241
QY 654 ArgGlyPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerIleValIleVal 673
Db 242 AGAAGTCTTCTCTCTCTTCAACATCCCTGCTTGATAGTAGCGCTAGCTGTGTT 301
QY 674 ValSerAsnLeuIleAspLeuGlyThrIleAlaIleAlaLeuLeuSerThrValSerVal 693
Db 302 TTAGTGAATGTCTGTGATGTCGGAACCATGTCACGCTGCTCAACGATCAGCGTC 361

QY 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
Db 362 ATGCTTATTTCTGCTTCTTCTGCTATGAGGGTTTGGGCTTATCCCAATATTCTTGGCG 421
QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733
Db 422 GAGATTTCCTCCACCTGCTGCTGCGGATCTGCATAGCATCTCGCGCTAACCTTCTGG 481
QY 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAlaIleGlyLeuAla 753
Db 482 ATCCGCGACATCATCTGATACATCTCCCGTGAAGCTCAATGCGATTGGTCTGCT 541
QY 754 GlyValPheSerIleTyrAlaValAlaCysLeuIleSerPheValPheValPheLeuIle 773
Db 542 GGAAGTCTGCGCATATATGCTGCTTGTGCTAGCTTGTATTTGCTTACATGAG 601
QY 774 ValProGluThrIleGlyMetProLeuGluValIleThrGluPhePheAlaValAla 793
Db 602 GTCCCTGAGACAAAGGGCATGCCCTGGAGGTATCATCCGAGTTCTTCTGTGCGGGCA 661
QY 794 LysGln 795
Db 662 AAGCAG 667

RESULT 6
US-09-291-922-5
; Sequence 5, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 870
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-291-922-5

Alignment Scores:
Pred. No.: 3,896-48 Length: 870
Score: 562.00 Matches: 103
Percent Similarity: 93.75% Conservative: 17
Best Local Similarity: 80.47% Mismatches: 8
Query Match: 13.79% Indels: 0
Gaps: 0
DB: 4

US-10-051-909-32 (1-800) x US-09-291-922-5 (1-870)

QY 669 LeuValIleLeuValAlaSerAsnLeuIleAspLeuGlyThrLeuAlaIleAlaLeuLeu 688
Db 8 GTTCTAACCTTGATTTGCTGCTCAATATTTCTGATGTGGGACCAATGCTCATGCTCATG 67
QY 689 SerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProIlePro 708
Db 68 TTCACAGTCAGTGCATCTTACTTCTGCTTCTTGTGATGGGGTTCGGCTTATTTCA 127
QY 709 AsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCys 728
Db 128 AACATTCTCTGTGAGAGATTTCCTCCAGACACCGTTGTGCGCATCTGATGCCATCTGT 187
QY 729 AlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsn 748
Db 188 GCCCTAACATCTGTGATGTCGATATCATTTGTGACATACACCTCCCGGTGATGCTCAAC 247

Qy	749	AlallelglyLeuAlaaglYValPheSerIleYrAlaValValCysLeuIleSerPheVal	768
Db	248	GCCATTGACCTCGCGAGAGTTTGGAATCTACGCGAGTGTCTGCATACCTACCGGCTTTCCTG	307
Qy	769	PheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhe	788
Db	308	TTTGTCTTCATCAAGGTGCGGAGACAAAGGCGATGCTCTTGAAATCATCAACCGAGTTC	367
Qy	789	PheAlaValGlyAlaIleGlnAla	796
Db	368	TTCTCTGTGGAGCAAGCAGGCC	391

RESULT 7

```

US-09-291-922-23
: Sequence 23, Application US/09291922
: Patent No. 6383776
:
: GENERAL INFORMATION:
: APPLICANT: Allen, Steve
: APPLICANT: Hitz, Bill
: APPLICANT: Kinney, Tony
: APPLICANT: Tiney, Scott
: TITLE OF INVENTION: Plant Sugar Transport Proteins
: FILE REFERENCE: BR-1163
:
: CURRENT APPLICATION NUMBER: US/09/291,922
: CURRENT FILING DATE: 1999-04-14
: EARLIER APPLICATION NUMBER: 60/083,044
: EARLIER FILING DATE: April 24, 1998
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: Microsoft Office 97
:
: SEQ ID NO: 23
:
: LENGTH: 1853
:
: TYPE: DNA
: ORGANISM: Glycine max
:
: US-09-291-922-23

```

Alignment Scores:

Pred. No.:	8, 71e-41	length:	1853
Score:	466.00	Matches:	165
Percent Similarity:	36.39%	Conservative:	109
Best local Similarity:	21.91%	Mismatches:	189
Query Match:	12.17%	Indels:	290
DB:	4.	Gaps:	18

US-10-051-909-32 (1-800) x US-09-291-922-23 (1-1853)

[illegible]

QY	197	sergllygmetpneuSerTyrCysmetValPheglymetSer---leuSerProser	215
Db	665	AATGAGGGATATTATTATGATACATCAACTATGCTTTCCAGCTGACACTAAG	724
QY	216	ProAspTPArgileMetleuGlyValleuAlaileProSerleuPhePheheglyleu	235
Db	725	GTCGGATGCCAATATGATCTTGAGATTGATGATCAATACCTTCGTA---CTCCTAACAGTA	781
QY	236	ThrilePheTyrleuProGluSerProArgTrpIleuValSerIleGlyArgMetAlaGlu	255
Db	782	GGAGTGTGGCGATGCCGAGATCCCAAGTCGCTTGATGAGGGGTGTTGGAGAG	841
QY	256	AlaIlySlyValleuGlnIlyleuArgGlyLysAspAspValSerGlyIleuSerleu	275
Db	842	GCAAGAAAAGTCTTAACAA-----	862
QY	276	LeuIeuGlnIlyleuGlnValGlyLysAspThrSerIleGlnIlyTrileileGlyPro	295
Db	862	-----	862
QY	296	AlaThrGlnAlaAlaAspAspIleuValThrAspGlyAspIlyGlnIleThrIleTyr	315
Db	863	-----ACCTCAGACAGCAAGAGAG-----	883
QY	316	GlyProGlnIlyGlnIlyGlnSerTrpIleAlaArgProSerIyGlyProIleMetleuGly	335
Db	883	-----	883
QY	336	SerValleuSerleuAlaSerArgHisGlySerMetValAsnGlnSerValProIleuMet	355
Db	884	GCCTCACTAAGCTTACCG-----	901
QY	356	AspProIleValThrleuPheGlySerValHisGluAsnMetProGlnAlaGlyIySer	375
Db	902	-----GAATCAACAACCCGAGGGATC-----	925
QY	376	MetArgSerThrleuPheProAsnDheGlySerMetPheSerValThrAspGlnHisAla	395
Db	926	CCCGAGACTTGC-----	937
QY	396	LysAsnGlnGlnTrpAspGlnGluAsnMetHisArgAspAspGlnIlyTrpAlaSerAsp	415
Db	938	-----AACGACGAC-----	946
QY	416	GlyAlaGlyIyAspTyrGlnAspAsnIleHisSerProIleuSerArgGlnAlaThr	435
Db	946	-----	946
QY	436	GlyAlaGlnIlyLysAspIleValHisIleGlyHisArgGlySerAlaIeuSerMetArg	455
Db	947	-----GTCGTTCAAGTAAT-----	961
QY	456	ArgGlnThrleuLeuGlyGlnIlyGlyAspGlyValSerSerThrAspIleGlyIyGly	475
Db	962	AAACAAAC-----AACGTAAGGTGA-----	985
QY	476	TrpGlnIleuAlaTrpIlySerGlnIlyLysGlnIyGlyLysGlnIyArgIyGlyIy	495
Db	985	-----	985
QY	516	SerIeuProGlnIlyGlyAspValPheGlnIySerGlnPheValHisAlaAlaIeu	535
Db	985	-----	985
QY	536	ValSerGlnSerAlaIeuPheSerIyGlyLysAlaGlnProArgMetSerAspAlaAla	555
Db	985	-----	985

```

QY 556 MetValHisProSerGluValAlaAlaValGlySerArgTrpIleAspLeuPhe----- 573
Db 986 -----TGGAAAGAGCTCTTCCTAT 1006
QY 574 ---GluProGlyValAlaArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln 592
Db 1007 CCACCGCCGCAATTCGACATCGAATCGCTGCCCTGGATTAATCTCTCCAAACA 1066
QY 593 PheAlaGlyIleAsnGlyLeuLeuTrpTrpTrpProGlnIleLeuGlnAlaGlyVal 612
Db 1067 GCGTCGGCGGTACAGCCGCTCTTGTGTACAGCCCAAGATCTTCGAAAGAGCTGGAT 1126
QY 613 AlaValIleLeuSerIlePheGlyLeuSerSerAlaSerIleLeuIleSerSer 632
Db 1127 ACAACGACGACGATTAAGCTT-----CTTGCAACGTCGGCCGTGGATTC 1171
QY 633 LeuTrpTrpLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetLeuSer 652
Db 1172 GTTAAGACCGGTCTCACTTGGCG-----GCTACGTTTACGTTGACCGCGTG 1219
QY 653 GlyArgArgPheLeuLeuLeuGlyTrpIleProIleLeuIleAlaSerLeuValIleLeu 672
Db 1220 GGTGTCGTCGCTGTATTTCTGTAGTGTGCGCGGACATGTGCTCTCTCCAGCTT 1279
QY 673 ValValSer---AsnLeuIleAspLeuGly-----ThrLeuAlaHisAlaLeuSer 689
Db 1280 GCGATCAGCTCAGCTGTATGTATCATTCGAGAGGAATTAATGCGCGCTGGATG 1339
QY 690 ThrValSer---ValIleValIleTrpPheCysCysPheValMetGlyPheGlyProIlePro 708
Db 1340 AGCATAGCCATGTGTGCTTACGTCGACGCGCATCTTCATCGTGGCGCTCCCATCAG 1399
QY 709 AsnIleLeuCysAlaGluIlePheProThrArgValArgIleuGlyLeuGlyIleAlaIleCys 728
Db 1400 TGGGCTATAGTTCGAGATCTTCCGTTGAGGCTGCGGCGGACAGTGGCGCGGGA 1459
QY 729 AlaPheTrpPheTrpIleGlyAspIleIleValIleTrpSerLeuProValMetLeuAsn 748
Db 1460 GTTGGCGGTAAATAGACACATGACGCGGTGTCTCATATGCTTCTGCTCCACTAGA 1519
QY 749 AlaIleGlyLeuAlaGlyValPheSerIleTrpAlaValAlaCysLeuIleSerPheVal 768
Db 1520 GCCATCAGCTATTGTGAGACTTCTTCTTATGTGTGCGATTCAGTGTGGGTGATA 1579
QY 769 PheValIlePheLeuValProGluTrpIleGlyMetProLeuGluValIleThr---Glu 787
Db 1580 TTCTTTTACACCGCTCTGCTTACAGACCGGGGAAAAAGCGCTCGAAGACATGGAAGGCT 1639
QY 788 PhePheAlaValGlyAlaValGlnAlaAlaValAla 800
Db 1640 TTTGGTACTTTTAAAGTCCAAATCCAAAGCCGACGAGGCT 1678

```

RESULT 8
US-09-291-922-21
Sequence 21, Application US/09291922
Patent No. 6383776

GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Hitz, Bill
APPLICANT: Kinney, Tony
APPLICANT: Tingey, Scott
TITLE OF INVENTION: Plant Sugar Transport Proteins
FILE REFERENCE: BB-1163
CURRENT APPLICATION NUMBER: US/09/291,922
EARLIER FILING DATE: 1999-04-14
EARLIER APPLICATION NUMBER: 60/083,044
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Microsoft Office 97
SEQ ID NO 21
LENGTH: 2017
TYPE: DNA
ORGANISM: Oryza sativa

```

US-09-291-922-21
Alignment Scores:
Pred. No.: 1,44e-40 Length: 2017
Score: 494.50 Matches: 185
Percent Similarity: 35.71% Conservative: 101
Best Local Similarity: 23.10% Mismatches: 206
Query Match: 12.13% Indels: 309
DB: 4 Gaps: 22
US-10-051-909-32 (1-800) x US-09-291-922-21 (1-2017)
QY 24 LeuProSerValIleLeuAlaLeuProGlyProLeuProAlaSerCysSerSerGln 43
Db 30 TTACACTGACCGGACATCTGATACAGGCCCA---GAGCAGGCTCTCTCTCTGCA 86
QY 44 GluProValThrSerAspAspIleLeuGluAsp-----LysMetSerGly--- 58
Db 87 CCACCGGAGATGGCTTCGCGCGCTCGGAGGCGCTCGCGCGGAGAGAGAGGCAAC 146
QY 59 -----AlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
Db 147 GTCCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 206
QY 76 AspAsnAlaThrIleAlaAlaAlaValLeuTrpIleGlyValGluPheGlnLeuGlnAsn 95
Db 207 GATATCGCGGAGATGAGCGGCGGCGCTGCTGATCATCAAGAGACTTCATCATC---AGT 263
QY 96 GluProThrValGlu-----GlyLeuIleValSerMetSerLeuIleGlyAlaThr 112
Db 264 GACGGAGAGTGGAGGTTCATGTGCGCATGCACTCACTGCTCATGCGCTCTTC 323
QY 113 IleValThrTrpPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIle 132
Db 324 GCG-----GCGGGCGGACGTCGATCGATCGGCGGCGGCGGATACCATCATCGT 371
QY 133 LeuSerSerIleLeuTrpPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTrp 152
Db 372 TTGCGCGCGCATATTCCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 431
QY 153 ValLeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValIleLeuVal 172
Db 432 ATGCTCATGTTCCGCGCGCTTCGTGCGCGCATCGCGCGGCGGCGGCGGCGGCGGCGG 491
QY 173 ProLeuTrpIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuPro 192
Db 492 CCGGCTACACCGCGCGAGGTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 551
QY 193 GlnPheSerGlySerGlyMetPheLeuSerTrpCysMetValPheGlyMetSerLeu 212
Db 552 GAGGTGTTCAATCACTTCGCGATCTGCTGCGGATGCTGCAATGCTTCTCCGCG 611
QY 213 SerPro---SerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhe 231
Db 612 TTGCGCGTGAACCTCGGCTGCGCATGCTCGGATCGGCGGCGGCGGCGGCGGCGGCGG 668
QY 232 PhePheGlyLeuTrpIlePheTrpLeuProGluSerProArgTrpLeuValSerGly 251
Db 669 CTGCTCGGCTCATGCTGCTCGGATCGCGAGTGGCGGCGGCGGCGGCGGCGGCGGCGG 728
QY 252 ArgMetAlaGluAlaValValLeuGlnValLeuArgGlyLysAspAspValSerGly 271
Db 729 CGGCTCGGAGCGGCGGAGGTGCTCGAAGAGCC----- 764
QY 272 GluLeuSerLeuLeuLeuGlyValGlyValGlyValGlyAspThrSerIleGluGluTrp 291
Db 765 -----TCCGACAG----- 773
QY 292 IleIleGlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspGluGln 311
Db 774 -----GCGAGAGAGCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 797
QY 312 IleThrLeuTrpGlyProGluGluGlyGlnSerTrpIleAlaArgProSerTrpGlyPro 331

```

```

Db 797 ----- 797
Qy 332 IleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSer 351
Db 797 ----- 797
Qy 352 ValProLeuMetAspProIleValThrLeuPheGlySerValHisGlyAsnMetProGln 371
Db 797 ----- 797
Qy 372 AlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThr 391
Db 797 ----- 797
Qy 392 AspGlnHisAlaLysAsnGlnGlnTrpAspGlnGlnAsnLeuHisArgAspAspGlnGln 411
Db 797 ----- 797
Qy 412 TyrAlaSerAspGlyAlaGlyGlyAspTyrGlnAspAsnLeuHisSerProLeuLeuSer 431
Db 797 ----- 797
Qy 432 ArgGlnAlaThrGlyAlaGlnGlyLysAspIleValHisHisGlyHisArgGlySerAla 451
Db 797 ----- 797
Qy 452 LeuSerMetArgArgGlnThrLeuLeuGlyGlyGlyAspGlyValSerSerThrAsp 471
Db 798 ----- 803
Qy 472 IleGlyGlyGlyTyrGlnLeuAlaTrpLysTyrSerGlyLysGlyGlyLysAsnGlyArg 491
Db 804 ATCAAGGCCGCC----- 815
Qy 492 LysGlnGlyGlyPheLysArgValTyrLeuHisGlnGlnGlyValProGlySerArgArg 511
Db 816 ----- 839
Qy 512 GlySerIleValSerLeuPro-----GlyGlyGlyAspValPheGlnGlySerGlnPhe 529
Db 840 GCGCAGCGTGCACCGCTCCCAAGAGAGAGAGCGGAAAC----- 878
Qy 530 ValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerGlyLeuAlaGlnPro 549
Db 879 ----- 884
Qy 550 ArgMetSerAspAlaAlaMetValHisProSerGlnValAlaAlaLysGlySerArgTyr 569
Db 885 CCGGTG----- 893
Qy 570 LysAspLeuPhe-----GlnProGlyValAlaArgAlaLeuLeuValGlyValGly 586
Db 894 AAGGAGCTCATCTCTCCCGCAGCCCGGCGACCGCGCATCTCTCTCCCGGAGTCGCG 953
Qy 587 IleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrThrProGlnIle 606
Db 954 ATCCACTTCTTCACATGCGCTGGCGATCTCCGCTGCTTCTTCAACGCCCTCTCGG 1013
Qy 607 LeuGlnGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAla 626
Db 1014 TTCACAGAGCCCGGATTAAAG-----AACGACAAACACTTCTGGGACACACTTGGCGG 1067
Qy 627 SerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMet 646
Db 1068 TTCGGTGTACACAGAGCTTTCATCTGTG-----GCGACT 1106
Qy 647 LeuLeuMetAspLeuSerGlyValArgPheLeuLeuLeuGlyThrIleProIleLeuIle 666
Db 1107 TTCCTCATGACGCGCTCGGCGCGCGCTGTGTGCTGGGACGACGCGGCGGATTAAAC 1166
Qy 667 AlaSerLeuVal-----IleLeuValIleSerAsnLeuIleAspLeuGly 681

```

```

Db 1167 CTCCTCCATCGGCTCGGCGCGGCGTCAACCGCTCGGCGGACGACCCCGGCGGAAAG 1226
Qy 682 ThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheVal 701
Db 1227 ATACCTTGGGCCATCGGCTTAAGCATGCGCTTCAACCTCGGCTACGTGCGCTTCTTCC 1286
Qy 702 MetGlyPheGlyProIleProAsnIleLeuCysAlaGlnIlePheProThrArgValArg 721
Db 1287 ATCGGCTTGGCCCCCATACGTGGGTGACAGCTCGGAGATCTTCCCGCTCAGGTGGC 1346
Qy 722 -----GlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleVal 739
Db 1347 GCGCTGGCTGCTCGCTCGGCTCGGCGCGCAACCGCGTCAACGAGCGGTATCTTCATG 1406
Qy 740 ThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleThr 759
Db 1407 ACCTTCTCG-----TCGCTTCCAGGCGCATCAACATCGGCGGAGGCTTCTTCTTAC 1460
Qy 760 AlaValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGly 779
Db 1461 TCCGCGCATCGCGCGGCTCGCTGGGTGTTCTTACACCTACCTCCGAGACCGCGCGC 1520
Qy 780 MetProLeuGlnValIleThrGlnPhePheAlaValAlaLysGlnAlaAlaLys 799
Db 1521 CGGACGCTGAGAGATGACCAAGCTGTC-----GGCGACAGCGCGCGCTCGGAA 1574
Qy 800 Ala 800
Db 1575 TCA 1577

RESULT 9
US-09-291-922-19
; Sequence 19, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291, 922
; EARLIER FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083, 044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Zea mays
US-09-291-922-19

Alignment Scores:
Pred. No.: 2,67e-40 Length: 1914
Score: 491.50 Matches: 172
Percent Similarity: 35.36% Conservative: 107
Best Local Similarity: 21.80% Mismatches: 201
Query Match: 12.06% Indels: 309
DB: 4 Gaps: 18

US-10-051-909-32 (1-800) x US-09-291-922-19 (1-1914)
Qy 41 SerSerGlnGlnProValThr-----SerAspAspIle-----LeuGlnAspLys 55
Db 35 TCAAGAAGTAGCGGTAAACGATGCTCCGACGAGCTCGCAAGCGCGTGCAGCCAGG 94
Qy 56 MetSerGly-----AlaValLeuValAlaIleValAlaSerIleGlyAsnLeu 71
Db 95 AAGAAAGGCAAGTCAGATATGCTCCATATATGCAATCTCGGCTCCATGAGCCCTGTTC 154
Qy 72 LeuGlnGlyTyrAspAsnAlaThrIleAlaAlaValLeuTyrIleLysGlnPhe 91

```

Db 155 ATCTTGCTATGACATGGGTGATGAGTGAAGCGGCATGATACATCAAGAAGCATCTG 214
 Qy 92 GlnLeuGlnAsnGlnProThrValGlnGlyLeuIleValSerMetSerLeuIleGlyAla 111
 Db 215 AATATC--ACGACCTGCACTGGAGATCTGATCGGATCTCATGCTCTAC----- 265
 Qy 112 ThrIleValThrThrPheSerGlyPro--LeuSerAspSerIleGlyArgArgProMet 130
 Db 266 TCGCTGTTCGATCTCTCGCTGGCGCGGAGCTCCAGACGATCGCGGCGCTTGAC 325
 Qy 131 LeuIleLeuSerSerIleLeuTyrrPhePheSerGlyLeuIleMetLeuTrpSerProan 150
 Db 326 GTCCGTTCGCGCTCATCTCTTCCTGCGGCTCATCTGCTCATGAGGTTCGCCGTC 385
 Qy 151 ValIleValLeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThr 170
 Db 386 TAGCGCATCTCATGCGCGCGCGCTTCGAGCGCGAGTGGGTGGGTACCGGGGCATG 445
 Qy 171 LeuValProLeuTyrrIleSerGlnIleAlaProSerGlnIleArgGlyLeuLeuAsnThr 190
 Db 446 ATCGCGCGCTGACACGCGCGAGATCTGCGCTGCGCGCTCCGCTGCTTACACAC 505
 Qy 191 LeuProGlnPheSerGlySerGlyGlyMetPheLeuSerTyrrCysMetValPheGlyMet 210
 Db 506 TTCGCGAGGTGTCATCAACATCGGCATCTGCTGCTACCTGCTCAACTTCGCGTTC 565
 Qy 211 SerLeuSerPro--SerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSer 229
 Db 566 GCGCGCGCTCCGCTCACTCGAGCTGGCGCGTCAATCTGCAATTCGCGCAATTCGCTCC 625
 Qy 230 LeuPhePhePheGlyLeuThrIlePheTyrrLeuProGlnSerProArgTrpLeuValSer 249
 Db 626 --GGCGTCTGCGCTCTCGGTTCCTGCATGCGCGAGTCCGCTCGGCTGCTGCTTG 682
 Qy 250 LysGlyArgMetAlaGlnAlaIleValValLeuGlnLysLeuArgGlyLysAspAspVal 269
 Db 683 AAGGCGCGCTGCGAGCGCGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 721
 Qy 270 SerGlyGlnLeuSerLeuLeuGlnGlyLeuGlnValGlyLysAspThrSerIleGln 289
 Db 721 ----- 721
 Qy 290 GlnTyrrIleIleGlyProAlaThrGlnAlaAlaAspAspLeuValThrAspGlyAspLys 309
 Db 721 ----- 721
 Qy 310 GlnGlnIleThrLeuTyrrGlyProGlnGlnGlnGlnSerTrpIleAlaArgProSerLys 329
 Db 722 -----ACCTTCGCCACGCGCAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 760
 Qy 330 GlnProIleMetLeuGlySerValLeuSerIleAlaSerArgHisGlySerMetValAsn 349
 Db 760 ----- 760
 Qy 350 GlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGlnAsnMet 369
 Db 760 ----- 760
 Qy 370 ProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSer 389
 Db 760 ----- 760
 Qy 390 ValThrArgGlnHisAlaLysAsnGlnGlnTrpArgGlnGlnAsnLeuHisArgAspAsp 409
 Db 760 ----- 760
 Qy 410 GlnGlnTyrrAlaSerAspGlyValGlyGlyAspTyrrGlnAspAsnLeuHisSerProLeu 429
 Db 760 ----- 760
 Qy 430 LeuSerArgGlnAlaThrGlyAlaGlnGlyLysAspIleValHisHisGlyHisArgGly 449
 Db 760 ----- 760

Qy 450 SerAlaLeuSerMetArgArgGlnThrLeuLeuGlnGlnGlyGlyAspGlyValSerSer 469
 Db 760 ----- 760
 Qy 470 ThrAspIleGlyGlyTyrrGlnLeuAlaTrpTyrrTrpSerGlyLysGlnGlyAsn 489
 Db 761 --GACATCAAGCGCGCG-- 775
 Qy 490 GlyArgLysGlnGlyGlyPheLysArgValTyrrLeuHisGlnGlnGlyValProGlySer 509
 Db 776 -----GGCGGATTCGCAAGCGGC 793
 Qy 510 ArgArgLysSerIleValSerLeuPro-----GlnGlyGlyAspValPheGly 525
 Db 794 CTCGACGGGAGACGTAAGTACGTAACCGGCAAGACGAGCGGGCGGTGAAGTTC----- 847
 Qy 526 GlySerGlnPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGly 545
 Db 847 ----- 847
 Qy 546 LeuAlaGlnProArgMetSerAspAlaAlaMetValHisProSerGlnValAlaAlaLys 565
 Db 847 ----- 847
 Qy 566 GlySerArgTrpLysAspLeuPhe-----GlnProGlyValArgArgAlaLeuLeu 582
 Db 848 --CAGGTGGAAGAAAGATATCTGCTCCGACCGCGCGGTGTCGACGATCTGCTC 904
 Qy 583 ValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrr 602
 Db 905 TCGCGCGGTGCTTCACCTTCTCCAGAGAGCTTCGAGACGACTCCGCTGCTCAGATC 964
 Qy 603 ThrProGlnIleLeuGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSer 622
 Db 965 AGCGCGCGCTGTTCAAGAGCGCGGAGTACACGACGACAAACAGCTCTCGGCGTCAAC 1024
 Qy 623 SerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuMetLeuProCysIle 642
 Db 1025 TGCCTG-----GTGGCGTACCAAGACGTTCTTCATCTCTG----- 1060
 Qy 643 GlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThrIle 662
 Db 1061 --GTGGCAGCTTCCTGCTGAGACGCGCGGGGTGGCGCTGCTGCTATACGACG 1117
 Qy 663 ProIleLeuIleAlaSerLeuValIleLeu-----ValValSerAsnLeu 677
 Db 1118 GCGCGAGTATGTCGCTGCTCATCTGCTCGGCTCGGAGGCTCACCGTCCGCGGACATCAC 1177
 Qy 678 IleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValIleValIleValTyrrPhe 697
 Db 1178 CCGACACCAAGGTCCGCGTGGCGCGCTGTCGATCCGCGTACACCTGCTTACATC 1237
 Qy 698 CysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGlnIlePhePro 717
 Db 1238 GCTTCCTTCTTCATCGGCTCGGCGCATACAGCGCGGTGTAACCTCGGAAATTTCCG 1297
 Qy 718 ThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIle 737
 Db 1298 CTCGAGGTGCGCGCTGCGCTGCGGAGTGGAGTGGCGAGAACCGCGTACACAGCGCC 1357
 Qy 738 IleValThrTyrrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSer 757
 Db 1358 GTCATCTCCATGACCTTCTGCTCCCTCCAAAGGCGCATACACATCGCGGAGGCTTTC 1417
 Qy 758 IleTyrrAlaValValCysLeuIleSerPheValPheValPheLeuLysValProGlyThr 777
 Db 1418 CTCATCTCCGACGCGCGCGCGGTGCTTGAGTTTCTTCTTACAGTCCCTCCGGAACA 1477
 Qy 778 LysGlyMetProLeuGlnValIleThrGlnPhePhe-----Ala 790
 Db 1478 CCGCGCGGACGCTGAGAGATGCGCAAGCTGTTCCGATGCGACAGCGCGGATGGCT 1537

QY 791 ValGlyAlaLysGlnAlaAlaLys 799
 DB 1538 GAAGAGCAGAGAGCCGCCAGCCAAAG 1564
 RESULT 10
 US-09-291-922-27
 ; Sequence 27, Application US/09291922
 ; Patent No. 6383776
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Hitz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Kinney, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB-1163
 ; CURRENT APPLICATION NUMBER: US/09/291,922
 ; EARLIER APPLICATION NUMBER: 60/083,044
 ; EARLIER FILING DATE: April 24, 1998
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 27
 ; LENGTH: 1872
 ; TYPE: DNA
 ; ORGANISM: Triticum aestivum
 US-09-291-922-27
 Alignment Scores:
 Pred. No.: 4.65e-40 Length: 1872
 Score: 489.00 Matches: 165
 Percent Similarity: 34.22% Conservative: 106
 Best Local Similarity: 20.83% Mismatches: 215
 Query Match: 12.00% Indels: 306
 DB: 4 Gaps: 14
 US-10-051-909-32 (1-800) x US-09-291-922-27 (1-1872)
 QY 24 LeuProSerValAlaLeuAlaLeuProGlyProLeuProProAlaSerCysSerSerGln 43
 DB 119 ATGGCTTTCCTGCGCTCCCGAGCCGGGAGCAGTCCATCCAGAAC----- 166
 QY 44 GluProValThrSerAspAspIleGluLysPheMetSerGlyAlaValLeuValAla 63
 DB 167 -----AAGGCAATTTCAAGTACGCTTCACCTGCGCC 199
 QY 64 IleValAlaSerIleGlyAsnLeuLeuGlnGlyTTPAspAsnAlaThrIleAlaAla 83
 DB 200 CTCTGCTTCCTCCATGCGCACCATCTCTCGGCTACGACGTTGGGATGAGCGGTGCG 259
 QY 84 ValLeuTyrIleLysLysGluPheGlnLeuGlnAsnGlu-----ProThrValGluGly 101
 DB 260 TCGCTGTCATCAAGAGGACCTGCAGATCAGCGACGTGAGCTGAGATCATGTGAGG 319
 QY 102 LeuIleValSerMetSerLeuIleGlyAlaThrIleValIleThrPheSerGlyProLeu 121
 DB 320 ATCCGAGCGGTACGCGCTCATCGGTCTTCCTC-----GGCGCAGAGAG 367
 QY 122 SerAspSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSer 141
 DB 368 TCCGCTGGGTGCGCGCGCTGCACCGTCTTCGCGCGCGCCATCTTCAACACAGCGC 427
 QY 142 GlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeuLeuAlaArgPheValAsp 161
 DB 428 TCTTGTCTGCTGCGCTTCGCGGTCACTAGCCATGCTCATGTGTGGGCGCTTCCTCAC 487
 QY 162 GlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaPro 181
 DB 488 GGAATCGGCGGTGCGGTACGCATCATGTGCGCGCAGTATACCCCGAGGTGTCCCG 547
 QY 182 SerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyGlyMetPhe 201
 DB 548 GCGTGGCGCGCGCTTCTCTCACGCTTTTCAACGAGGTGTTTCATCATGTGGGCTCTC 607

QY 202 LeuSerTyrCysMetValPheGlyMetSerLeuSerPro---SerProAspTyrPargIle 220
 DB 608 CTTCGCTACGCTCCCACTACGCTTCGCGGCTCCCGCTCCACCTCAGCTGCGGCTC 667
 QY 221 MetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeu 240
 DB 668 ATGCTCGCAATCGCGCGCTCCCGCTCCGCGC---CTGCTTGCGCTCAAGGTTCGCGCATG 724
 QY 241 ProGluSerProArgTyrPheValSerLysGlyArgMetAlaGluAlaLysValLeu 260
 DB 725 CCGAGTCTCTCTGCTGCTGCTCATAGAAAGCCGCTCGGAGCGCCAGGCGCGCTCTG 784
 QY 261 GluLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeuLeuGluGlyLeu 280
 DB 784 ----- 784
 QY 281 GluValGlyLysPheThrSerIleGluGluTyrIleIleGlyProAlaThrGluAla 300
 DB 784 ----- 784
 QY 301 AspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGly 320
 DB 784 ----- 784
 QY 321 GlnSerTyrIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeu 340
 DB 784 ----- 784
 QY 341 AlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThr 360
 DB 784 ----- 784
 QY 361 LeuPheGlySerValHisGluAsnMetProGluAlaGlyGlySerMetArgSerThrLeu 380
 DB 784 ----- 784
 QY 381 PheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTyr 400
 DB 785 -----GCCAAGACTCCGACACG 802
 QY 401 AspGluGlnAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyLysAsp 420
 DB 803 CCGAGAGAGCGCGGAGCGCTTGACGATCAAGAGCTCC----- 844
 QY 421 TyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLys 440
 DB 844 ----- 844
 QY 441 AspIleValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu 460
 DB 844 ----- 844
 QY 461 GlyGluGlyLysAspGlyValSerSerThrAspIleGlyGlyGlyTyrGlnLeuAlaTyr 480
 DB 844 ----- 844
 QY 481 LysTyrPheSerGluLysGluGlyLysAsnGlyArgLysGluGlyGlyPheLysArgValTyr 500
 DB 844 ----- 844
 QY 501 LeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGly 520
 DB 845 -----GCCGCAATCCCTAGGAACTTGCAGGACGAGTGTGCTGATCTTACG--- 892
 QY 521 GlyAspValPheGluGlySerGluPheValHisAlaAlaValSerGlnSerAla 540
 DB 892 ----- 892
 QY 541 LeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSer 560
 DB 893 -----ACAAAGCGCGCCAGAGAACAGCAGTG----- 919
 QY 561 GluValAlaAlaLysGlySerArgTyrLysAspLeuPhe-----GluProGlyVal 577

```

Db      920 -----TGAAGAGACTCATCTTTTCCGCCAGCCCAAGCCATG 955
Qy      578 ARGAGALALEuLeuValGlyIleGlnIleLeuGlnIlePheAlaGlyIleLeu 597
Db      956 CGCGCGATACCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1015
Qy      598 GlyValLeuTyrrThrProGlnIleLeuGlnIleGlnIleValAlaValIleLeuSer 617
Db      1016 TCCGTCGCTCTATAGCCACCGCTTCCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1075
Qy      618 LysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrrThrLeuLeu 637
Db      1076 CTGCTCGCG-----GCCACATGGCGCATGCGGGGTATAGAAAGCGCTCTTC 1120
Qy      638 MetLeuProCysAlIleGlyPheAlaMetLeuMetAspLeuSerGlyArgArgPheLeu 657
Db      1121 ATCCGTG-----GTGGCCACGTTCCAGCTCAGCGCGCGCGCGCGCGCGCGCGCG 1168
Qy      658 LeuLeuGlyThrIleProIleLeuIleAlaSerLeuVal-----IleLeu 672
Db      1169 CTGCTGACAGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1228
Qy      673 ValValSerAsnLeuIleAspLeuGlyThrLeuAlaIleAlaLeuLeuSerThrrValSer 692
Db      1229 GTCGAGGTGCGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1288
Qy      693 ValIleValIrrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCys 712
Db      1289 ATCTTGCGCTTACGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1348
Qy      713 AlaGlnIlePheProThrrArgValArgGlyLeuCysIleAlaIleCysAlaIlePheThrrPhe 732
Db      1349 TCGGAGGTCTTCCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1408
Qy      733 TrpIleGlyAspIleIleValIleThrrTyrrSerIleuProValMetLeuAsnAlaIleGlyLeu 752
Db      1409 CGCGTCACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1468
Qy      753 AlaGlyValPheSerIleTyrrAlaValValCysLeuIleSerPheValPheValPheLeu 772
Db      1469 GCGCGCAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1528
Qy      773 LysValProGluThrrIleGlyMetProLeuGluValIleThrrGluPhePhe----- 789
Db      1529 TTCAATTCGCGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1588
Qy      790 -----AlaValGlyAlaLysGlnAlaAlaAlaLys 799
Db      1589 GACACGCGCGCTCGAAGCCCAAGACACCGCGCGCGCGCGCGCGCGCGCGCGCG 1624

RESULT 11
US-09-291-922-25
; Sequence 25, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Kinney, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 25
; LENGTH: 2089
; TYPE: DNA
; ORGANISM: Trifolium aestivum
US-09-291-922-25

```

```

Alignment Scores:
Pred. No.: 2,34e-37 Length: 2089
Score: 463.50 Matches: 166
Percent Similarity: 35.82% Conservative: 108
Best Local Similarity: 21.70% Mismatches: 200
Query Match: 11.37% Indels: 291
DB: 4 Gaps: 17

US-10-051-909-32 (1-800) x US-09-291-922-25 (1-2089)
Qy      52 LeuGluAspLysMetSerGly-----AlaValLeuValAlaIleValAlaSer 67
Db      104 GTGAGGCCCAAGAAAGAGCGCAAGTACGTTCCGCTTCGCGCGCGCGCGCGCGCGCGCG 163
Qy      68 IleGlyAsnLeuLeuGlnIleTyrrAspAsnAlaThrrIleAlaAlaValLeuTyrrIle 87
Db      164 ATGACCTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 223
Qy      88 LysIleGlyPheGlnIleLeuGlnAsnGluProThrrValGlu-----GlyLeuIleVal 104
Db      224 CAGAAAGATCTGAAGATC--AACGACACCCAGCTGAGGCTCTCATAGGCGCATCTCAAC 280
Qy      105 SerMetSerLeuIleGlyAlaThrrIleValThrrPheSerGlyProLeuSerAspSer 124
Db      281 GTGATCTGCTCATTTGGCTCTTCCGCG-----GCCGGGCGGAGCTCGGACATCG 328
Qy      125 IleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrrPhePheSerGlyLeuIle 144
Db      329 ATCGCGCGCGCGCTTCCACCATCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 388
Qy      145 MetLeuTrpSerProAsnValTyrrValLeuLeuLeuAlaArgPheValAspGlyPheGly 164
Db      389 ATGGCTTCTCTCTCAATCACTAGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 448
Qy      165 IleGlyLeuAlaIleThrrLeuValProLeuTyrrIleSerGlnIleAlaProSerGlnIle 184
Db      449 GTGGGAGTACGCTTCATGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 508
Qy      185 ArgGlyLeuLeuAsnThrrLeuProGlnPheSerGlySerGlyMetPheLeuSerTyrr 204
Db      509 CGTGGGCTTCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 568
Qy      205 CysMetValPheGlyMetSer---LeuSerProSerProAspTrpArgIleMetLeuGly 223
Db      569 GTCTTCACTTGGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 628
Qy      224 ValLeuAlaIleProSerLeuPhePheGlyLeuThrrIlePheTyrrLeuProGluSer 243
Db      629 ATAGGCGCGGTGCGCGCTCGTC---CTGCTGCGTTCAATGCGTCCGCGCATGCGCGAG 685
Qy      244 ProArgTrpLeuValSerLysGlyArgMetAlaGlnAlaLysLysValLeuGlnLysLeu 263
Db      686 CCGCGGTGCTCGTCATGAAGGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 745
Qy      264 ArgGlyLysAspAspValSerGlyGluLeuSerLeuLeuGlnIleGlyValGly 283
Db      745 ----- 745
Qy      284 GlyAspThrSerIleGluGluTyrrIleIleGlyProAlaThrrGlnAlaAspAspLeu 303
Db      746 TCCGACACG----- 754
Qy      304 ValThrrAspGlyAspLysGluGlnIleThrrLeuTyrrGlyProGluGluGlnSerTrp 323
Db      755 -----CCGGAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 775
Qy      324 IleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerAlaSerArg 343
Db      776 ATCGCC----- 781
Qy      344 HisGlySerMetValaGlnIleSerValProLeuMetAspProIleValThrrLeuPheGly 363

```

Db	781	-----	-----	-----	781
Qy	364	SerValHisGluAsnMetProGlnAlaGlySerMetArgSerThrLeuPheProAsn			383
Db	781	-----	-----	-----	781
Qy	384	PheGlySerMetPheSerValThrAspGlnHisAlaIleAsnGlnGlnTrpAspGlnGlu			403
Db	781	-----	-----	-----	781
Qy	404	AsnLeuHisArgAspAspGluGluTrpAlaSerAspGlyAlaGlyIleAspTyrGluAsp			423
Db	782	-----	-----	-----	782
Qy	424	AsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLeuAspIleVal			443
Db	799	-----	-----	-----	799
Qy	444	HisHisGlyHisArgGlySerAlaLeuSerMetArgGlnGlnThrLeuGlnGlyGluGly			463
Db	799	-----	-----	-----	799
Qy	464	GlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSer			483
Db	799	-----	-----	-----	799
Qy	484	GluIysGlnGluGluAsnGlyArgLysGlnGlyGlyPheLysArgValTyrLeuHisGln			503
Db	799	-----	-----	-----	799
Qy	504	GluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyIleAspVal			523
Db	800	-----	-----	-----	800
Qy	524	PheGluGlySerGluPheValHisAlaIleAlaLeuValSerGlnSerAlaLeuPheSer			543
Db	845	-----	-----	-----	845
Qy	544	LysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAla			563
Db	848	AAAGGAGCAGCAGAGAGAGAGCGGCTTTGAAAGCAGCTCAGCTGTACCGACCATACCC			907
Qy	564	AlaIysGlySerArgTyrTrpLysAspLeuPheGluProGlyValArgArgAlaLeuVal			583
Db	908	-----	-----	-----	908
Qy	584	GlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrThr			603
Db	929	GGAATCGGCATCCTCTTCACGAGCTTCGCGGATCGAGCCGTCGTCTACACG			988
Qy	604	ProGlnIleLeuGlnGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSer			623
Db	989	CCGCTAGTTTCAAGAGCGCGCGCATACG			1030
Qy	624	---AlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIle			642
Db	1031	CTCCGCGGCACACCGCTGGCGGTCCGGCGCACCAATACGTCCTTATCATCTG			1081
Qy	643	GlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIle			662
Db	1082	---GTGGCACCTTCTCTCTGACCGCATCCGCGCGCGCGCTGATGCTTACACGACG			1138
Qy	663	ProIleLeuIleAlaSerLeuVal			677
Db	1139	GCGCGCAGCTCGTCTCTTACGTGGCGCTCGGACGCGGGCTACCGCTATACGCGCAC			1198
Qy	678	IleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPhe			697
Db	1199	CCGGACGAGAAAGATCACCTGGCGCATCGTCTGTCATCTTCATCATGAGCCTAACGTG			1258
Qy	698	CysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGlnIlePhePro			717
Db	1259	GCGTTCTTCTCCATCGGCTCGGCGCCCATCAGTGGGTGATACGGCTGGAATCTTCCCG			1318

```

QY 718 thrtatgvalaaggllyaucyslllealllecyalaphethrphetrpilleglyaspile 737
Db 1319 CTGACGCTGCGCGCGCTGGCTGCTCTCCCTGGCGCTGCGCTCAACCGCTGACCGCGC 1378
QY 738 lilevalthrtYrSerleuProvalaMctleuabnaAllelglyleualaglyvalPheser 757
Db 1379 GTGATCTTCATGACCTTCATTGGCTGTGTCCAAAGCCATGACCATGACCGGGCGCTTCTTC 1438
QY 758 lileTrralvalaValcYsleuilleserPhevalaPhevalaPheleuYsValProgluThr 777
Db 1439 CTCTTCGCGCGGACATGCGCTCATTCGACATGGGTGTCTTCTTCGCTCACTGCGCGAGACC 1498
QY 778 lyeGlymetProleuagluValallethrtGluPhePhealvalaGlyla-----lyeGln 795
Db 1499 GCGGCGCGGACCGCTGAGAGACATGACTCGCTTTCGGCAACAGCGCCACGACACAGCAG 1558
QY 796 AlaAlaAlaAlaAla 800
Db 1559 GCGCGCGCGGAGACC 1573

RESULT 12
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.: 2.59e-26 Length: 4411529
Score: 412.00 Matches: 162
Percent Similarity: 33.46% Conservative: 111
Best Local Similarity: 19.85% Mismatches: 204
Query Match: 10.11% Indels: 340
DB: 3 Gaps: 17

US-10-051-909-32 (1-800) x US-09-103-840A-1 (1-4411529)
QY 2 ArgserglyserTrpleuAlaValGlnThrPro-----PheThrPro 15
Db 3717093 CGCGTGTGGCAGCCCTCATGCAAAATACACAGCCCTATACGACGCGGGGTTTCAGCCC 3717152
QY 16 -----AppLeuaspArgGluArgleuLeuProSerValValleuAla 30
Db 3717153 GCCCGGCGCGCCGACGACCCCGACCGC-----GTGGTGGACGTG 3717191
QY 31 leuProGlyProleuPro-----ProAlaSerCysSerSerGlnGluProvalThr 47
Db 3717192 CTGAGCGCGCGCGGTCACTGACCA----- 3717221
QY 48 seraspAspIleleuGluaspIlysmetSerGlyAlaValleuValAlaIleValAlaSer 67
Db 3717222 -----GATCGGGGGCGCGGGCG-CTCTCTGTCGGGCTCACCGCGCC 3717262
QY 68 lileGlyleuLeuGlnGlyTTPAspAsnAlaThrIleAlaAlaAlaValleuYrIle 87
Db 3717263 AGCGTCGCGCTCTTACGGGTACGACCTTTCGCGCATGCGGGGTGGTGGTCTCTCTC 3717322

```

Oy	88	yslysglnpheglnleuglnamsluprothrtValgluglyleuilevalsermetSer	107
Db	3717323	AGCGAGAAATTCGAATCTCAACCACTCGAGAA-----CAAGATCTCAACCAACCGCGG	3717378
Oy	108	leuilegialathrtilevalthrThrPheSerGlyProLeuSerAspSerilegIyArg	127
Db	3717377	GTGCTGGCCAGATAGCGCGGCGCTTGCGGCGCGGATCTCTCCAAACCGATCGGAGCG	3717436
Oy	128	ArgProMetLeuileleuSerSerileleuTyrrPheSerGlyLeuileMetLeuTrp	147
Db	3717437	AAGAAATCGGTGCTCTCATCTGCGCGGATACGCAAGTTCCTCCCTCTCGGCGGACG	3717496
Oy	148	SerProAsnValValleuLeuLeuValaArgPheValAspGlyPheGlyilegIyLeu	167
Db	3717497	TGGGTGTCGATACGATGCTGGTGGTGCGGCTCTGCTGGGTGTGATCAATCGGCTG	3717556
Oy	168	AlaValThrLeuValProleuTyrrIleSerGluilealProSerGluileArgIyLeu	187
Db	3717557	TCCGTGTGTGTGTGTCGCGATGTATGTGCGCGAGTCCGCGCGCGCGCTGTGGGTCTG	3717616
Oy	188	LeuAsnThrLeuProGlnPheSerGlySerGlyIyMetPheLeuSerTyrrCysMetVal	207
Db	3717617	TTGTGATCCGCGTATCACTGACGTGCGACGCTTACCGGATCTGTCTGTAC-----CTG	3717670
Oy	208	PheGlyMetSerLeuSerProSerProAsnTrpArgIyleMetLeuGlyValleuValle	227
Db	3717671	GTCGGCTACCTGTTGTGCGGATCGCACGGCTGCGCGAGTTCGCGGCTGTGGCTGCGCGCG	3717730
Oy	228	ProSerLeuPhePhePheGlyleuThrIlePheTyrrLeuProGlnSerProArgTrpLeu	247
Db	3717731	CCGGCGACCGCTGCTGTCCCGTTG---TTGTGCGCGATGCCGAGTACCGCCCGCTGGTAT	3717787
Oy	248	ValSerIlysgIyArgMetAlaGluValaLysIlyValleuGlnIlyLeuValArgGlyLysAsp	267
Db	3717788	CTGCTCAAGGGCGCGATCGCGACGCGGTAGCGGCTGTGCGCGGATCCACCGGAGGCC	3717844
Oy	268	AspValSerGlyGlnLeuSerSerLeuLeuGlnGlyLeuGluValaGlyIyAspThrSer	287
Db	3717848	GACATCGATGCCGAGCTGAGCC-----	3717866
Oy	288	IleGluGluTyrrIleilegIyProAlaThrGluValaAspAspLeuValThrAspGly	307
Db	3717868	-----	3717866
Oy	308	AspIlysgIuGlnIleThrLeuTyrrIyProGluGlnGlyGlnSerTrpIleAlaArgPro	327
Db	3717868	-----	3717866
Oy	328	SerIlysgIyProIleMetLeuGlySerValleuSerIleuAlaSerArgHisGlySerMet	347
Db	3717868	-----	3717866
Oy	348	ValAsnGlnSerValProleuMetAspProIleValThrLeuPheGlySerValHisGlu	367
Db	3717868	-----	3717866
Oy	368	AsnMetProGlnAlaGlyIlySerMetArgSerThrLeuPheProAsnPheGlySerMet	387
Db	3717868	-----	3717866
Oy	388	PheSerValThrAspGlnIleAlaLysAsnGluGlnTrpArgGlnGluAsnLeuHisArg	407
Db	3717868	-----	3717866
Oy	408	AspAspGlnGluTyrrAlaSerAspGlyAlaGlyIyAspTyrrGluAspAsnLeuHisSer	427
Db	3717869	---GATATGCGCGCGCGGATCGACAGACGCGCGCGGTATCGCGAATGATGTGCGCGCG	3717925
Oy	428	ProLeuLeuSerArgGlnAlaThrIyAlaGluGlyLysAspIleValHisHisGlyHis	447
Db	3717926	CCGATATCTG-----	3717934

QY	448	ArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGlnGlyAlaAspGlyVal	467
Db	3717934	-----	3717934
QY	468	SerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpIlyStrpSerGluLeuGly	487
Db	3717934	-----	3717934
QY	488	GluAsnGlyArgGlyGlyGlyGlyPheLeuArgValTyrLeuHisGlnGlnGlyValPro	507
Db	3717934	-----	3717934
QY	508	GlySerArgArgGlySerIleValSerLeuProGlyGlyAspValPheGlnGlySer	527
Db	3717934	-----	3717934
QY	528	GluPheValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAla	547
Db	3717934	-----	3717934
QY	548	GluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySer	567
Db	3717934	-----	3717934
QY	568	ArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyAlaIle	587
Db	3717935	-----CGGGCCAGCGCTTCCATCCAGCGCTC	3717961
QY	588	GlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeu	607
Db	3717962	GGCTTCTCGTCCGATACCGGGATACACCGCATCTTCTACATCTACGTCCGACCTTTC	3718021
QY	608	GluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerAspAlaSer	627
Db	3718022	GGCCGATGGGCTTCGGG-----GGCATTTCCGATGCTTGGCCCTGCGCGATG	3718072
QY	628	IleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeu	647
Db	3718073	GTGCAGTGGCGGCTTGGCGCGCGGTG-----TGTGCC-----TTCCTGTTT	3718114
QY	648	LeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAla	667
Db	3718115	CTGGTCATCGGCTGGGCGGCTTCCGCCCATCTTGTTCGCCGATCGGACGATATATACC	3718174
QY	668	-----SerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHis	685
Db	3718175	GCAGATGCGGTCGATACCGTATTCGCCACAACATCCGATGGTGACAG-----GGG	3718228
QY	686	AlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGly	705
Db	3718229	CTGGTGTGGGGTTCGGCGGCGTCTCTGTTCATTCAGGTGGTCACTTCGATTCGGC	3718288
QY	706	ProIleProAsnIleLeuCysAlaGlnIlePheProThrArgValArgGlyLeu---Cys	724
Db	3718289	TGCGTGGTCTGGGGTACCGCGCGGAGAGCTTCCGTCGCGGCGGTGATGGGATCG	3718348
QY	725	IleAlaIleCysAlaPheThrPheTrpIleGlyAspIleValIleThrTyrSerLeuPro	744
Db	3718349	AGCCGATGCTGCACCTGCACACTACGCGCAACGATCTGTGCGGCTTCTGCTC---	3718405
QY	745	ValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValAlaCysLeu	764
Db	3718406	ACCATGTGCTGCGTTCGCGGCGGACAGGCGTTTCCCGGCTTCGCGACGTTGCGCGTC	3718465
QY	765	IleSerPheValPheValPheLeuLysValProGlnThrLysGlyMetProLeuGlnVal	784
Db	3718466	GTCCGCTTCTGGTCTGCTGACCGCTTTCGCGGAGACCAAGGCGCCCAACTCGAGGAG	3718525
QY	785	IleThrGluPhePheAlaValAlaGlyAlaIleGlnAlaIleAlaLysAla	800
Db	3718526	ATCCGGGACCTTCTGGGAGAACGCGCGCGCTTCGCGCCCGCCAGCGGTCA	3718573


```

Db      833 CAACGAGAGAACCCGGCCAAAGTGTCTAAAGAGCTCGCGGAGACGTGACGTAC 892
Qy      270 rGlyGluLeuSerLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluG 290
Db      893 CCATGACCTG----- 902
Qy      290 uTyrlleIleGlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspGlyG 310
Db      902 ----- 902
Qy      310 uGlnIleThrLeuTyrlGlyProGluGluGlyGlnSerThrIleAlaArgProSerLeG 330
Db      902 ----- 902
Qy      330 yProIleuLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnG 350
Db      902 ----- 902
Qy      350 nSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetPr 370
Db      902 ----- 902
Qy      370 oGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerVa 390
Db      902 ----- 902
Qy      390 lThrAspGlnHisAlaGlyAsnGluGlnThrAspGluGluAsnLeuHisArgAspAspG 410
Db      903 -----CAGAGATGAGAGAGAGAG- 920
Qy      410 uGluTyrlAlaSerAspGlyAlaGlyGlyAspTyrlGluAspAsnLeuHisSerProLeu 430
Db      920 ----- 920
Qy      430 uSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySe 450
Db      921 -AGTGGGAG----- 929
Qy      450 rAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyAspGlyValSerSerTh 470
Db      930 -ATGATGGCGGAGAGAGAGAGTCCATCTCTG----- 959
Qy      470 rAspIleGlyGlyTyrlGlnLeuAlaThrTyrlSerGlnGlyGluGluAsnG 490
Db      959 ----- 959
Qy      490 yArgGlyGluGlyGlyPheTyrlArgValTyrlLeuHisGlnGluGlyValProGlySerAr 510
Db      959 ----- 959
Qy      510 gArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheVa 530
Db      959 ----- 959
Qy      530 lHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSerIleGlyLeuAlaGluProAr 550
Db      959 ----- 959
Qy      550 gMetSerAspAlaAlaMetValHisProSerGluValAlaAlaGlySerArgTyrlPly 570
Db      959 ----- 959
Qy      570 sAspLeuPhe--GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnI 589
Db      960 -GAGCTGTTCGCTCCCGGCTACCGGCGGAGCCCATCTCCATCGCTGTGTCTCAGCT 1018
Qy      589 eLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrlTyrlThrProGlnIleLeuGluG 609
Db      1019 GTCCGAGCAGCTGTCTGTGATCAAGCTGTCTTCTATTACTCAGAGCATTTTGAGAA 1078
Qy      609 nAlaGlyVal--AlaValIleLeuSerIlePheGlyLeuSerSerAlaSerAlaSerI 628
Db      1079 GCGCGGGGGGAGCAGAGCTGTGTATGCCACATTGGCTCGGTATTCGTTCAACAGCGCTT 1138

```

```

Qy      628 eLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeu 648
Db      1139 CACTGTCTGTCTGCTG-----TTTGT 1159
Qy      648 uMetAspLeuSerGlyArgArgPheLeu--LeuLeuGlyThrIleProIleLeu--Il 666
Db      1160 GTTGAGGAGGAGGAGCGCGGAGCCTTGACCTTATAGGCTCGCTGGCATGGGGGTG 1219
Qy      666 eAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAl 686
Db      1220 TGCCATCTCATGACCATCGCGCTAGCAGCTGTGAGACAGCTACCTCGAGTCTCTAT- 1277
Qy      686 aLeuLeuSerThrValSerValIleValTyrlPheCysCysPheValMetGlyPheGlyPr 706
Db      1278 ----CTAGCATCGTGGCCATCTTTGGCTTTGTGGCTTTCTTAAGTGGATCTGGGCC 1333
Qy      706 oIleProAsnIleLeuGlyCysAlaGluIlePheProThrArgValArgGlyLeuGlyAl 726
Db      1334 CATCCATGTTTCATCGTGGCTGAATCTTACGACGAGGTCACAGCTGCATGCTGC 1393
Qy      726 aIleCysAlaPheThrPheThrIleGlyAspIleIleValThrTyrlSerLeuProValMe 746
Db      1394 CTTTGACAGCTTCTTCCAACTGACCTCAAAATTTCATTGTGGGACATGCTTCCAGTATGT 1453
Qy      746 tLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrlAlaValAlaCysLeuIleSe 766
Db      1454 GGAGCAACTGTGTGT--CCCTACGTTTATCATCTTCACTGTGCTCGTGTCTGTT 1510
Qy      766 rPheValPheValPheLeuGlyValProGluThrGlyGlyMetProLeuGluValIleTh 786
Db      1511 CTTGATCTTCACTTCACTTCAAAAGTTCCTGAGACCTAAAGCCGAGACCTTCATGATGCC 1570
Qy      786 rGluPhePheAlaValGlyAlaGlyAlaGlyAlaAlaAlaLys 799
Db      1571 TTCCGGCTTCCGGGAGGAGGAGCCAGCCAAAGTGAATAG 1610

RESULT 15
US-09-480-884A-135
; Sequence 135; Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongrong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 135
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-480-884A-135

Alignment Scores:
Pred. No.: 5,676-26 Length: 2856
Score: 355.00 Matches: 164
Percent Similarity: 33.29% Conservative: 107
Best Local Similarity: 20.15% Mismatches: 206
Query Match: 8.71% Indels: 337
DB: 4 Gaps: 23

US-10-051-909-32 (1-800) x US-09-480-884A-135 (1-2856)
Qy      18 AspArgArgGluArgLeuLeuProSerValValLeuAlaLeuProGly- 33
Db      86 GACCGAGACAGAGCTGTAGGAGGAGAGCGCGCTCGGACGCCCGTGCACCGCGTAC 145

```

QY 34 -----ProleuProProAla-SerCySerSerGlnGluProValThrSerAspAs 50
 DB 146 CCGGGGCGACGACAGGACACCGGACG---GCTGCCATGAGGCCGACGACGACG----- 197
 QY 50 pIleuGluuAapLyMetSerGlyAlaValLeuValAlaIle---ValAlaSerIleG 69
 DB 198 -----AAGCTGACGGGTGCTCATGCTGCTGCTGAGGAGGACGAGCTTGG 244
 QY 69 yAsnLeuLeuGlnGlyTrpAspAsnAlaThrIleAlaAlaValLeuTyrlleGly 89
 DB 245 CTCCTGCGAGTTGGCTTACACACCTGAGATCAATGCCCCCGACAGGTGATC---GA 301
 QY 89 sGluPheGlnLeuGln-----AsnG 96
 DB 302 GGAAGTTCTACACACGACATGGGTCCACCGCTATGGGAGACATCTGCCACACGCT 361
 QY 96 uProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrTh 116
 DB 362 CACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421
 QY 116 rPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSerI 136
 DB 422 CTCCTGGGCTTTTCTGTTAACCGCTTTGGCCGGGGAATTCATGCTGATGATGACCT 481
 QY 136 eLeuTyrrPhePheSerGlyLeuIleMetLeuTrpSer-----ProAsnValTyrrVa 153
 DB 482 GCTGGCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 541
 QY 153 lLeuLeuLeuAlaArgPheValaAspGlyPheGlyIleGlyLeuAlaValThrLeuValPr 173
 DB 542 GCTGATCTCTGGGCGCTTCATCATCGTGTACTGCGGCTGACACACAGGCTTGTGCTG 601
 QY 173 oLeuTyrrIleSerGlnIleAlaProSerGlnIleArgGlyLeuLeuAsnThrLeuProG 193
 DB 602 CATGATGATGGGTGAAGTGTACCCACACGCTTTGTTGGGCGCTTGGGACACCTGCACCA 661
 QY 193 nPheSerGlySerGlyMetPheLeuSerTyrrCySerMetValPheGlyMet---Serie 212
 DB 662 GCTGGGCGATCGTCGTCGCGATCCTCATCGCC-----CAGTGTCTCGGCTGATCCTCAT 715
 QY 212 uSerProSerProAsp---TrpArgIleMetLeuGlnValLeuAlaIleProSerLeuPh 231
 DB 716 CATGGGCAACAAGACGCTGTGGCCCTGCTGAGCATCATCTTCCGCGGCTGCT 775
 QY 231 ePhePheGlyLeuThrIlePheTyrrLeuProGluSerProArgTrp---LeuValSerIy 250
 DB 776 GCAAGTGC---ATCGTGTGCTGCCCTTCTGCGCCGAGAGTCCCGCTTCTGCTCATCAACCG 832
 QY 250 sGlyArgMetAlaGluAlaIleValLeuGlnIleValArgGlyLyAspAspValSe 270
 DB 833 CACACGAGAGAACCGGGCCACAGAGTGTCTTAAGAAGCTGCCGAGACAGCTGAGTAC 892
 QY 270 rGlyGluLeuSerLeuLeuLeuGlnGlyLeuGlnValaGlyGlyAspThrSerIleGlu 290
 DB 893 CCATACCTG----- 902
 QY 290 uTyrrIleIleGlyProAlaThrGlnAlaAlaAspAspLeuValThrAspGlyAspLyG 310
 DB 902 ----- 902
 QY 310 uGlnIleThrLeuTyrrGlyProGluGluGlnIleSerTrpIleAlaArgProSerLyG 330
 DB 902 ----- 902
 QY 330 yProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValaAsnG 350
 DB 902 ----- 902
 QY 350 nSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetPr 370
 DB 902 ----- 902

QY 370 oGlnAlaGlyLySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerVa 390
 DB 902 ----- 902
 QY 390 lThrAspGlnHisAlaIleValAsnGlnGlnTrpAspGlnGlnIleAsnLeuHisArgAspAspG 410
 DB 903 -----CAGAGAGATGAAGAGAGAG----- 920
 QY 410 uGlnTyrrLaserAspGlyAlaGlyGlyAspTyrrGluAspAsnLeuHisSerProLeuLe 430
 DB 920 ----- 920
 QY 430 uSerArgGlnAlaThrGlyAlaGlnGlyLyAspIleValHisHisGlnHisArgGlySe 450
 DB 921 -AGTCGGAG----- 929
 QY 450 rAlaLeuSerMetArgArgGlnThrLeuLeuGlnGlyGlyGlyAspGlyValaSerSerTh 470
 DB 930 -ATGATGGGAGAGAAAGGTCCACCATCTG----- 959
 QY 470 rAspIleGlyGlyTyrrTrpGlnLeuAlaTrpLySerTrpSerGlnuLyGlnGlnGlnAlaAsnG 490
 DB 959 ----- 959
 QY 490 yArgLyGlnGlyGlyPheLyAsrGlyTyrrLeuHisGlnGlnGlyValaProGlySerAr 510
 DB 959 ----- 959
 QY 510 gArgGlySerIleValSerLeuProGlyGlyGlyAspValaPheGlnGlySerGluPheVa 530
 DB 959 ----- 959
 QY 530 lHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLyGlyLeuAlaGluProAr 550
 DB 959 ----- 959
 QY 550 gMetSerAspAlaAlaMetValHisProSerGlnuValaAlaAlaLyGlySerArgTrpLy 570
 DB 959 ----- 959
 QY 570 sAspLeuPhe---GluProGlyValaArgArgAlaLeuLeuValaGlyIleGlnI 589
 DB 960 -GAGCTGTTCCGCTCCCGGCTTACCGGACGACCCATCTTCCCTGTGGTGTGCTGAGCT 1018
 QY 589 eLeuGlnGlnPheAlaGlyIleAsnGlyValaLeuTyrrTyrrThrProGlnIleLeuGln 609
 DB 1019 GTCCAGAGCTGTGCTGACCAAGCTGTCTTCTATTACTCCACGACATCTTGAGAA 1078
 QY 609 nAlaGlyVal---AlaValIleLeuSerTyrrPheGlyLeuSerSerAlaSerAlaSerI 628
 DB 1079 GCGCGGGGTGACAGACGCTGTATGCCACCATTTGGCTCCGCTATCGTCAACACGGGCTT 1138
 QY 628 eLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLe 648
 DB 1139 CACTGTCTGTGCTG-----TTTGT 1159
 QY 648 uMetAspLeuSerGlyArgArgPheLeu---LeuLeuGlyThrIleProIleLeu---Il 666
 DB 1160 GTGGAGCGACGACGCGCGGAGCTTGCACCTTATGACCTCGCTGACATGGCGGGGTTG 1219
 QY 666 eAlaSerLeuValIleLeuValaIleSerAsnLeuIleAspLeuGlyThrLeuAlaHisAl 686
 DB 1220 TGGCATATCTATGACATCCGCGACGACGTGAGACAGGTACCCGATGTCAT-- 1277
 QY 686 aLeuLeuSerThrValSerValIleValTyrrPheCySerPheValMetGlyPheGlyPr 706
 DB 1278 -----CTGAGCATCTGCGCATCTTGGCTTGTGGCTTCTTTCAGAGTGGGTCTCGGCC 1333
 QY 706 oIleProAsnIleLeuCyAlaGlnIlePheProThrArgValaArgGlyLeuCyIleAl 726
 DB 1334 CATCCCATGTTCAATCGTGTGAATCTTTCAGGACGAGGTCACAGCTGACCTGCATTCG 1393
 QY 726 aIleCyAlaPheThrPheTrpIleGlyAspIleIleValaThrTyrrSerLeuProValMe 746


```
Db      1394 CGTTGCAGGCTTCTCCAACTGACCTCAAAATTTCATGTGGGCGATGCTTCCAGTATGT 1453
Oy      746 cLeuSmlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSe 766
Db      1454 GGAGCACTGTGTGGT---CCCTACGCTTTCATCACTTCACTGTGTCTGTCTGTGT 1510
Oy      766 rPheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleTh 786
Db      1511 CTTGATCTTCACTTCAAAAGTTCTGAGACTAAAGCCGACCTTGAATGAGATCGC 1570
Oy      786 rGluPhePheAlaValGlyAlaLysGlnAlaAlaLys 799
Db      1571 TTCCGGCTTCCGGCAGGGGAGCCAGCCAAAGTATAG 1610
```

Search completed: January 6, 2004, 01:58:24
Job time : 12167 secs

THIS PAGE BLANK (USPTO)

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed

Result No.	Score	Query Match	Length	DB	ID	Description
1	3517	86.3	2824	24	ABK51962	Corn contig encodi
2	3517	86.3	2824	25	ABX93198	CDNA encoding corn
3	2674	65.6	2601	24	ABK51965	Soybean contig enc
4	2674	65.6	2601	25	ABX93201	CDNA encoding soyb
5	2594	63.4	2190	24	ABZ13510	Arabidopsis thalia
6	2275	53.8	2205	24	ABZ14449	Arabidopsis thalia
7	1670.5	41.0	1692	24	ABK51966	Soybean cDNA clone
8	1670.5	41.0	1692	25	ABX93202	CDNA encoding soyb
9	1424	34.9	1487	24	ABK51968	Wheat contig encod
10	1424	34.9	1487	25	ABX93204	CDNA encoding whea
11	943	23.1	1009	24	ABK51969	Wheat cDNA clone w
12	943	23.1	1009	25	ABX93205	CDNA encoding whea
13	587.5	14.4	1518	24	ABL41880	Nucleotide sequenc
14	562	13.8	870	24	ABK51964	Rice cDNA clone r1
15	562	13.8	870	25	ABX93200	CDNA encoding rice
16	539	13.2	1395	24	ABK7616	Bacillus lichenifo
17	520.5	12.8	1826	21	AAQ42332	Arabidopsis thalia
18	496	12.2	1853	24	ABK51971	Soybean contig enc
19	496	12.2	1853	25	ABX93209	CDNA encoding soyb
20	494.5	12.1	2017	24	ABK51972	Rice contig encodi
21	494.5	12.1	2017	25	ABX93208	CDNA encoding rice
22	493.5	12.1	1527	21	AAAC4261	Arabidopsis thalia
23	493	12.1	1482	24	ABZ12990	Arabidopsis thalia
24	491.5	12.1	1914	24	ABK51971	Corn cDNA clone ce
25	491.5	12.1	1914	25	ABX93207	CDNA encoding corn
26	489	12.0	1566	24	ABZ13964	Arabidopsis thalia
27	489	12.0	1872	24	ABK51975	Wheat cDNA clone w
28	489	12.0	1872	25	ABX93211	CDNA encoding whea
29	488	12.0	1925	24	ABL41874	Nucleotide sequenc
30	488	12.0	2056	24	ABL41873	Nucleotide sequenc
31	481	11.8	1766	24	ABL41872	Nucleotide sequenc
32	475	11.7	1644	24	ABZ12515	Arabidopsis thalia
33	468	11.5	1662	20	AAZ24474	Spinach glucose tr
34	467	11.5	2134	21	AAAC36954	Arabidopsis thalia
35	465.5	11.4	1690	24	ABL41879	Nucleotide sequenc
C 36	463.5	11.4	1662	24	ABL41878	Nucleotide sequenc
37	463.5	11.4	2089	25	ABK51974	Wheat cDNA clone w
38	463.5	11.4	2089	24	ABX93210	CDNA encoding whea
39	453.5	11.1	1905	22	AAH88793	Sugar transporter
40	453.5	11.1	3370	19	AAV11474	T. halophilus xyliu
41	452	11.1	12354	22	AAAG46243	DNA encoding novel
42	452	11.1	12354	22	AAAG46261	DNA encoding novel
43	447.5	11.0	349980	22	AAH61967	C glutamicum codin
44	447.5	11.0	349980	22	AAH61966	C glutamicum codin
45	444	10.9	1473	22	AAH65168	C glutamicum codin

RESULT 1
ABK51962

ID	Accession	Location/Qualifiers
XX	ABK51962 standard; cDNA; 2824 BP.	
AC	ABK51962;	
XX	27-AUG-2002 (first entry)	
DT		
XX	Corn contig encoding Arabidopsis thaliana-like sugar transport protein	
DE		
XX	Corn; Arabidopsis thaliana-like sugar transport protein;	
KW	carbohydrate transport; grain filling; annual field crop;	
KW	plant; gene; ss.	
XX		
OS	Zea mays.	
XX		
FH	Key	

FT CDS 238..2481
 FT /*tag= a
 FT /product= "Corn Arabidopsis thaliana-like sugar
 FT transport protein"
 XX US6183776-B1.
 XX
 XX 07-MAY-2002
 XX
 XX 14-APR-1999 99US-0291922.
 XX
 XX 24-APR-1998 98US-083044P.
 XX
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 XX Allen SM, Hitz WD, Kinney AJ, Tingey SV;
 XX
 XX WPI, 2002-451386/48.
 XX
 XX P-Psdb; AAU97201.
 XX
 PT New nucleic acid encoding plant sugar-transport proteins, useful for
 PT preparing transgenic plants with altered carbohydrate distribution
 XX
 XX Claim 3; Column 25-28; 54pp; English.
 XX
 CC The present invention relates to the isolation of plant polynucleotide
 CC sequences encoding an Arabidopsis thaliana-like sugar transport
 CC protein or Beta vulgaris-like sugar transport protein. The
 CC polynucleotide sequences are useful for altering the level of sugar
 CC transport proteins in plants, i.e. for control of carbohydrate transport
 CC and distribution in plant cells, e.g. during grain filling of annual
 CC field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
 CC carbohydrate flows and sugar transport. The polynucleotide
 CC sequences can also be used to isolate cDNA sequences and genes that
 CC encode homologues of the new proteins. The present sequence
 CC representing a contig assembled from various corn cDNA clones encodes
 CC an Arabidopsis thaliana-like sugar transport protein.
 XX
 XX Sequence 2824 BP; 644 A; 649 C; 745 G; 775 T; 11 other;
 XX
 Alignment Scores:
 Pred. No.: 3,816-298 Length: 2824
 Score: 3517.00 Matches: 684
 Percent Similarity: 95.18% Conservative: 27
 Best Local Similarity: 91.57% Mismatches: 34
 Query Match: 86.31% Indels: 2
 DB: Gaps: 2
 US-10-051-909-32 (1-800) x ABK51962 (1-2824)
 QY 56 MetSerGlyAlaValLeuValAlaValAlaSerIleGlyAsnLeuLeuGlnGlyTyr 75
 DB 238 ATGGGGGGCGCGATGATGCTCCCATCGCGGCTCTATCGGCACTTGTGAGGAGCTGG 297
 QY 76 AspAsnAlaThrIleAlaAlaValAlaValLeuTyrIleValGlyGlnLeuGlnAsn 95
 DB 298 GAGCATGGGACATTCCTGACACCGCTCTCTCATTAAGAGAGATTCACCTGACAGAC 357
 QY 96 GluProThrValGlyGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
 DB 358 GAGCCTCGATCGAAGGCTCATCTCGCATGTTCTCATATGGGCAACAGTCACTACA 417
 QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyAspArgProMetLeuIleLeuSer 135
 DB 418 ACATCTCCGGGGCCAGGGCTGACCTGCGTGTAGAGAGCCCATCTGCTGCTGCTGCT 477
 QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeu 155
 DB 478 GTCTCTACTTCGTAGAGGGCTGGTGAATGCTTTGGGCGCCAAATGTGTACTCTTGGCTC 537
 QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
 DB 538 CTCGAAAGGCTCATGTGATGGTTCGGTATCGGTTTGGGGGTCACTTGTCTCTTAC 597

QY 176 IleSerGluIleAlaProSerGluIleArg---GlyLeuLeuAsnThrLeuProGlnPhe 194
 DB 598 ATCTCCGAAGTGCACCGCACAGANATTTCTGGGCTGNTNGAACACCTTCCGACATTC 657
 QY 195 SerGly---SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213
 DB 658 ATTGGCGGTCAAGNGAGGATGTTCTCTCTACTGATGATGCTTTGGGATGTCCTCATG 717
 QY 214 ProSerProAspTTPAglIleMetLeuGlyValLeuAlaIleProSerLeuPhePhe 233
 DB 718 CCCAACTGATTTGAGGCTCATGCTGAGTTCTGTCATCCGCTCATATATATCTT 777
 QY 234 GlyLeuThrIlePheTyrLeuProGlyLeuSerProArgTTPleuValSerGlyArgMet 253
 DB 778 GAACTGACTGCTTCTACTTCTCTGATTCACCAAGTGGCTTTGAGCAAGAGATG 837
 QY 254 AlaGluAlaValGlyValLeuGlnLeuValArgGlyValAspAspValSerGlyGluLeu 273
 DB 838 GCGAGGCGGAAGAGAGTGTNTCCAAAGGCTCGGGGAAGAGATGTCANGGAGANG 897
 QY 274 SerLeuLeuLeuGlyGlyLeuGlyValGlyValAspThrSerIleGlyGlyTyrIle 293
 DB 898 GCTCTTCTAGTTGAAGGTTTGGGCTCGTAAAGATACGTAATTNAGACTATCATATT 957
 QY 294 GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspGlyGlnIleThr 313
 DB 958 GAACTGCTCACCGAGGAGCGCATGATCTTGTACTGACGTTGATTAAGCAACATCA 1017
 QY 314 LeuTyrGlyProGluGluGlyGlnSerTyrIleAlaArgProSerGlyProIleMet 333
 DB 1018 CTTATAGGCGCTGAAGAAAGCCAGTCATGATTCGACCTTTAAGGAGCCCATCATG 1077
 QY 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
 DB 1078 CTGGAAGTGTCTTCTCTTCTGTCATCTCGATGAGCAAGTGAACCAAGTGTACCC 1137
 QY 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
 DB 1138 CTTATGATCCGATTTGACCTTTTGTGTGATGTCATGATGATGCTTCAAGCTGGA 1197
 QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
 DB 1198 GGAAGTATGAGAGACACATTTCTTCCAAACTTGGAGTATGTCAGTGTACAGATCAG 1257
 QY 394 HisAlaValAsnGluGlnTTPAspGlyGluAsnLeuHisArgAspAspGlyGlyTyrAla 413
 DB 1258 CATGCCAAMAATGACAGTGGATGAAGAGATCTTCAATAGGATGACGAGAGATACGCA 1317
 QY 414 SerAspGlyAlaGlyValAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
 DB 1318 TCTATGTGTCAGAGAGTGAATATGAGACATCTCCATAGCCCATTTGCTGTCCAGGCG 1377
 QY 434 AlaThrGlyAlaGluGlyValAspIleValHisHisGlyHisArgGlySerAlaLeuSer 453
 DB 1378 GCAACAGGTGGGAAGGAGAGACATTTGTGCACCATGTCACGTTGGAAGTCTTGAGC 1437
 QY 454 MetArgArgGlnThrLeuLeuGlyGlyGlyValAspGlyValAspSerThrAspIleGly 473
 DB 1438 ATGGAAGGCAAAACCTTTAAGGAGGAGTGAATGTTGACACACATGATATCGCT 1497
 QY 474 GlyGlyTyrGlnLeuAlaTTPLeuTyrSerGlyValAspGlyGluValArgGlyVal 493
 DB 1498 GGGGAGATGGACCTTGTGTAATGTCACAGAAAGAGGTGAATGTGTAAGAGAA 1557
 QY 494 GlyGlyPheValArgValTyrLeuHisGlnGluValProGlySerArgArgGlySer 513
 DB 1558 GGTGTTTCAAAAGATCTACTTGCACCAAGAGGAGATTCCTGCTCAAGAAAGGGGCTCA 1617
 QY 514 IleValSerLeuProGlyGlyValAspValPheGluGlySerGluPheValHisAlaAla 533
 DB 1618 ATTGTTTCACTTCCCGGTGGTGGAGATGTTCTTGAAGGTAGTGTATGTATCAGTCTGCT 1677

QY 534 AAlaLeuValSerGlnSerAlaLeuPheSerIysGlyLeuAlaGluProArgMetSerAsp 553
 DB 1678 GCTTAGTAGAGTCAGTCAGCACTTTCTCAAGAGGCTCTTGAGACCAAGCATGTGCAAT 1737
 QY 554 AAlaMetValHisProSerGluValAlaAlaIysGlySerArgTyrPheAspLeuPhe 573
 DB 1738 GCTGCATGAGTTCACCCCATCTGAGGAGTGTGCAAGAGTTCACGTTGGAAGATTGTTT 1797
 QY 574 GluProGlyValArgArgAlaLeuLeuValGlyValAlaGlyIleGlnIleLeuGlnIlePhe 593
 DB 1798 GAACCTGAGTAGGCGGTGCTGTAGTGGGTGTGGAATTCAGATCTTCAACAGTTT 1857
 QY 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGlnAlaGlyValAla 613
 DB 1858 GCTGGAATTAACGGTCTCTGATCATACCCCAAAATTTCTTGACCAAGCTGTGGA 1917
 QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
 DB 1918 GTTATCTTTCCAATTGGTCTCAGCTCGCATCAGATCCATCTTGATCAGTTCTCTC 1977
 QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
 DB 1978 ACTACCTTACTATCTCTCTCTGCAATGGCTTGGCCATGGCTTATGGATCTTCCGA 2037
 QY 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
 DB 2038 AGAAGCTTTTGTCTGCTGAGGACCAATTCATCTTGATGATCTCTGATCTTCTGTT 2097
 QY 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
 DB 2098 GTGTCAATCTATGATTTGGGTACACTAGCCCATGCTTCTCTCCACATCAGTGT 2157
 QY 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
 DB 2158 ATCGTCTACTTGTCTGCTCTGTATGAGGATTTGGTCCCATCCCAACATTTATGTGCA 2217
 QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTyr 733
 DB 2218 GAGATCTTCCACACAGGGTGTGTGGCTCTGTATGTCATTTGTGCTTTCATTTCTG 2277
 QY 734 IleGlyAspIleIleValIleThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
 DB 2278 ATCGAGATATCATGTCATCCTACACAGCTTCTCTGATGATGCTAAAGCTATGTGACTGG 2337
 QY 754 GlyValPheSerIleTyrAlaValAlaCysLeuIleSerPheValPheValPheLeuLys 773
 DB 2338 GGTGTTTTCAGCATATATGCACTGTATGCTGATTTCTTGTGTCTCTCTCTTAG 2397
 QY 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
 DB 2398 GTCCCTGAGACAAAGGGATGCCCTTGAGGTTATTAACGAATTTCTTGCAAGTTGGTGG 2457
 QY 794 LysGlnAlaAlaAlaLysAla 800
 DB 2458 AAGCAAGCGGCTGCAAAAGCC 2478
 RESULT 2
 ABX93198 standard; cDNA; 2824 BP.
 ID ABX93198;
 AC ABX93198;
 XX 29-MAY-2003 (first entry)
 DT 29-MAY-2003 (first entry)
 DE cDNA encoding corn sugar transport protein #1.
 KW Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
 KW plant sugar transport protein; carbohydrate transport; soybean;
 KW carbohydrate distribution; plant; gene; ss.
 OS Zea mays.
 XX
 PN US2002178468-A1.

XX 28-NOV-2002.
 PD 17-JAN-2002; 2002US-0051902.
 XX 24-APR-1998; 98US-083044P.
 PF 14-APR-1999; 99US-0291922.
 XX (ALLEN) ALLEN S. M.
 PA (HITZ) HITZ W. D.
 PA (KINNEY) KINNEY A. J.
 PA (TINGEY) TINGEY S. V.
 XX Allen SM, Hiltz WD, Kinney AJ, Tingey SV;
 DR WPI, 2003-340957/32.
 DR P-PSDB; ABD08326.
 PT Novel plant sugar transport proteins and nucleic acid encoding the
 PT protein useful for producing transgenic plants having altered levels of
 PT sugar transport protein
 PS Claim 2; Page 13-15; 56pp; English.
 XX The present invention relates to the isolation of Arabidopsis
 CC thaliana-like or Beta vulgaris-like sugar transport proteins, and
 CC the polynucleotide sequences encoding them. The plant sugar transport
 CC proteins of the invention have been isolated from corn, rice, soybean,
 CC and wheat. The polypeptides of the invention may be used for altering
 CC the level of expression of a sugar transport protein in a host cell,
 CC by transforming a host cell with a chimeric construct encoding all,
 CC or a portion of the sugar transport protein, in sense or antisense
 CC orientation. Particularly, the polypeptides may provide a means to
 CC control carbohydrate transport and distribution in plants.
 CC ABX93198-ABX93205 represent cDNA sequences encoding Arabidopsis
 CC thaliana-like sugar transport proteins.
 SQ Sequence 2824 BP; 644 A; 649 C; 745 G; 775 T; 11 other;
 Alignment Scores:
 Pred. No.: 3,81e-298 Length: 2824
 Score: 3517.00 Matches: 684
 Percent Similarity: 95.18% Conservative: 27
 Best Local Similarity: 91.57% Mismatches: 34
 Query Match: 86.31% Indels: 2
 DB: 25 Gaps: 2
 US-10-051-909-32 (1-800) x ABX93198 (1-2824)
 QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuGlnGlyTyr 75
 DB 238 ATGGGGGGCGCGTATGTCGTCATCGCGGCTCTATCGCAACTTCTGAGGGCTGG 297
 QY 76 AspaEnaIatnrlleAlaAlaValLeuTyrIleLysGluPheGlnLeuGlnAsn 95
 DB 298 GACATGTGACATATGCTGAGGCCCTCTGTACATAAAGAGATTTCACTGACAGAGC 357
 QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
 DB 358 GAGCCTCTGATCGAAGGCTCATCGTCATGTTCTCATATGGGGCAACAGTCATCA 417
 QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
 DB 418 ACATCTCCGGGCGCAAGGCTGACTGCTGTGTAGAGAGCCCATCTGTGCGCTGGCT 477
 QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu 155
 DB 478 GTCCCTTACTTCTGTCAGGGCTGTGATGCTTTGGGGCCAAATGTGTATCATCTTGTCTC 537
 QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
 DB 538 CTCGCAAGGCTCATGTATGAGGTTCGATTCGTTTGGGGGTACACACTTGTCTCTTAC 597

QY 116 IISerGluIleAlaProSerGluIleArg---GlyIleuLeuAsnThrLeuProGlnPhe 194
 DB 598 ATCTCGAAACCTGACACCGCAGANAAATCTTGGGGCTGNTGNAACACGTGGCCGAGTTC 657
 QY 195 SerGly---SerGlyGlyMetPheLeuSerTyrosMetValPheGlyMetSerLeuSer 213
 DB 658 ATTTGGGGTCAGAGGAGGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 717
 QY 214 ProSerProAspThrArgIleMetLeuGlyValIleuAlaIleProSerIleuPhePhe 233
 DB 718 CCCAAACCTGATGGAGGCTCATGCTTGAGATTCGTGATCCGCTCACTTATMTACTTT 777
 QY 234 GlyLeuThrIlePhePheTyrlleuProGlnSerProArgTyrIleuValSerIleGlyArgMet 253
 DB 778 GGAGCTGACTGCTTCTTACTTCTGATCATCCAAAGTGGCTTGTNAGCAAAAGAAAGATG 837
 QY 254 AlaGluAlaIleGlyValIleuGlnIleuLeuArgGlyIleuAspAspValSerGlyGluLeu 273
 DB 838 GCGGAGGCGAGAGAGTGTGCAAAAGCTGCGGAGAGAGATGTCTCANGGAGGANG 897
 QY 274 SerIleuLeuLeuGluGlyIleuGluValGlyIleuAspThrSerIleGluIleuTyrIleIle 293
 DB 898 GCTGCTTCTAGTTGAAGTGTGGGGGTGGTAAGATACAGTATTTNAGAGTACATCAT 957
 QY 294 GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspIleGluIleThr 313
 DB 958 GGAGCTGGCACCGAGGACGCGATATCTTGTAACTGACGCGATGAGAAACAATCACA 1017
 QY 314 LeuTyrGlyProGluGluGlyIleuSerTyrIleAlaArgProSerIleGlyProIleMet 333
 DB 1018 CTTTATGGGCGCTGAAAGAGGCGACATGATGCTTCGACCTTCAAGGAGCCCATCATG 1077
 QY 334 LeuGlySerValIleuSerIleuAlaSerArgHisGlySerMetValAsnGlnSerAlaPro 353
 DB 1078 CTTGGAAGTGTCTTCTCTGATCTGTCATGGAGATGTGTAACAGAGTGTACCC 1137
 QY 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAspMetProGlnAlaGly 373
 DB 1138 CTTAAGGATCCGATTTGTGACATTTTGTGTAGTGTCCATGAGAAATATGCTCAACTGGA 1197
 QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
 DB 1198 GGAATGTATGAGAGACATGTTGTTCCAACTTGGAGATGTTCAGTGTCAACATCAG 1257
 QY 394 HisAlaIleAsnGluIleThrAspGluIleuLeuHisArgAspAsnGluIleTyrAla 413
 DB 1258 CATGCCAAATAAGACACAGTGGATGAGAAATCTTCATAGGATGACAGAGATACCA 1317
 QY 414 SerAspGlyAlaGlyIleuAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
 DB 1318 TCTGATGGTGCAGAGGTGACATATGAGGACATCTCCATAGCCCATTTGCTGCCAGCGAG 1377
 QY 434 AlaThrGlyAlaGluGlyIleuAspIleValHisHisGlyIleuArgGlySerAlaLeuSer 453
 DB 1378 GCACACAGGTGCGGAGAGGAGACATGTGCAACATGTGCAACCGTGGAGAGTCTTTGAGC 1437
 QY 454 MetArgAspGlnThrLeuLeuGlyIleuGlyIleuAspAlaValSerSerThrAspIleGly 473
 DB 1438 ATGAGAAAGGCAAAAGCTCTTATAGGAGGAGTGGAGATGGGTAGAGCACTGATATTCGAT 1497
 QY 474 GlyIleTyrGlnIleuAlaTyrIleuSerGlyIleuGluGlyIleuAsnGlyIleuArgIleGlu 493
 DB 1498 GGGGGATGGAGCTTGTCTTGAATGGTTCAGAGAAAGGAGGAGATGTGTAAGAAAGAA 1557
 QY 494 GlyIlePheIleuArgValIleuHisGlnGluIleuValProGlySerArgArgGlySer 513
 DB 1558 GGTGGTTTCAAAAGAGTCTACTTGCACCAAGAGGAGTTCCTGCTCAAGAAAGGGGCTCA 1617
 QY 514 IleValSerLeuProGlyIleuGlyIleuAspValPheGluGlySerGluPheValHisAlaIle 533
 DB 1618 ATTGTTTCACTTCCGAGTGGTGGCATGTTCTTGAAGGAGTGAAGTGTGTAATCATCTGCT 1677
 QY 534 AlaIleuValSerGlnSerAlaIleuPheSerIleuAlaGluProArgMetSerAsp 553

DB 1678 GCTTATGATAGTCACTGACGACTTTTCCAAAGGGTCTGTGTAACACGATGTCAGAT 1737
 QY 554 AlaAlaMetValHisProSerGluValAlaAlaIleGlySerArgTyrIleuAspLeuPhe 573
 DB 1738 GCTGCAATGTTTCAACCAATCTGAGAGTGTGCAAAAGGTTTCACTTGAAGAAATTTGTT 1797
 QY 574 GluProGlyValArgArgAlaLeuLeuValGlyIleuGlyIleuGlnIleuGlnIleuPhe 593
 DB 1798 GAACCTGGAGTGAAGCGGTGCTTGTAGTCCGTTGGATTCAGATCTTCAACAGATT 1857
 QY 594 AlaGlyIleAsnGlyValIleuTyrTyrThrProGlnIleuGluGlnAlaGlyValAla 613
 DB 1858 GCTGGAATTAACCGGTGTTCTGTACTATATACCCCAAAATTTTGAGCAAGCTGGTGCA 1917
 QY 614 ValIleuSerIleuPheGlyIleuSerSerAlaSerAlaSerIleuIleuIleuSerIleu 633
 DB 1918 GTTATTTCTTCCAAATTTGCTCAGCTCGGCAATCAGATCCATCTTGATCAGTCTCTC 1977
 QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
 DB 1978 ACTATCTTACTATAGCTTCTCTGATTTGCTTTCATGCTGCTTATGATCTTTCCGGA 2037
 QY 654 ArgArgPheLeuLeuGlyThrIleProIleuIleuAlaSerLeuValIleuVal 673
 DB 2038 AGAAGTTTGTGCTGCTAGGACCAATTCATCTTGATAGCATCTTATATCTGTT 2097
 QY 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
 DB 2098 GTGTCCATCTTAATTTGATTTGGTATACCTACCTCATGCTTGTCTCCACCATAGTGT 2157
 QY 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
 DB 2158 ATGCTTACTTCTGCTGCTTGTATGGAGTTTGTGCTCATCCCAATTTTATGTGCA 2217
 QY 714 GluIlePheProThrArgValArgGlyIleuCysIleAlaIleCysAlaPheThrPheTyr 733
 DB 2218 GAGATCTTTCACACAGGGTGTGGCTGTGTATGTCATTTGCTTATACATTTGCG 2277
 QY 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAlaAlaIleGlyLeuAla 753
 DB 2278 ATCGAGATATCATCTGCACCTACAGCTTCTGTGATGTGCAATGTATTTGAGCTGCG 2337
 QY 754 GlyValPheSerIleTyrAlaValAlaCysLeuIleSerPheValPheValPheLeuIle 773
 DB 2338 GGTGTTTTCACATATATGACGTGATGCTTGTATTTCTTGTGTTGCTTCTTATAG 2397
 QY 774 ValProGluThrIleGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
 DB 2398 GTCCCTGAGACAAAGGGGATGCCCTTGAGGTTATTAACGAATCTTTGACAGTTGTCG 2457
 QY 794 LysGlnAlaAlaAlaIleVal 800
 DB 2458 AAGCAAGGCGCTGCMAAAGCC 2478
 RESULT 3
 ABKS1965
 ID ABKS1965 standard; cDNA; 2601 BP.
 AC ABKS1965;
 DT 27-AUG-2002 (first entry)
 DE Soybean contig encoding A. thaliana-like sugar transport protein.
 KW Soybean; Arabidopsis thaliana-like sugar transport protein;
 KW carbohydrate transport; grain filling; annual field crop;
 KW plant; gene; ss.
 OS Glycine max.
 OS
 FH Key Location/Qualifiers
 FT CDS 175..2388

/*tag= a
/product= "Soybean Arabidopsis thaliana-like sugar
transport protein"

US6383776-B1.

07-MAY-2002.

14-APR-1999; 99US-0291922.

24-APR-1998; 98US-083044P.

(DUPO) DU PONT DE NEMOURS & CO E I.

Allen SM, Hiltz WD, Kinney AJ, Tingey SV;

WPI; 2002-451386/48.

P-PSDB; MAU97204.

New nucleic acid encoding plant sugar-transport proteins, useful for
preparing transgenic plants with altered carbohydrate distribution -

Claim 3; Column 35-38; 54pp; English.

The present invention relates to the isolation of plant polynucleotide
sequences encoding an Arabidopsis thaliana-like sugar transport
protein or Beta vulgaris-like sugar transport protein. The
polynucleotide sequences are useful for altering the level of sugar
transport proteins in plants, i.e. for control of carbohydrate transport
and distribution in plant cells, e.g. during grain filling of annual
field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
carbohydrate flows and sugar transport. The polynucleotide
sequences can also be used to isolate cDNA sequences and genes that
encode homologues of the new proteins. The present sequence
representing a contig assembled from various soybean cDNA clones
encodes an Arabidopsis thaliana-like sugar transport protein.

Sequence 2601 BP; 648 A; 543 C; 644 G; 766 T; 0 other;

Alignment Scores:

Pred. No.: 2,946-224 Length: 2601
Score: 2674.00 Matches: 522
Percent Similarity: 82.20% Conservative: 92
Best Local Similarity: 69.88% Mismatches: 115
Query Match: 65.62% Indels: 18
DB: 24 Gaps: 8

US-10-051-909-32 (1-800) x ABK51965 (1-2601)

QY 56 MetSerGIYAlaValLeuValAlaValAlaSerIleGIYAsnLeuLeuGlnGlyTyr 75
DB 175 ATGAAGAGTGGCGCTTGTGCTAATGCGCGCTTCATTGTAATTCCTCCAGAGATGG 234
QY 76 AspaEnAlaThrIleAlaAlaValLeuTyrIleValGluPheGlnLeuGlnAsn 95
DB 225 GATATGCTACATGCGCGGCTAATGCTTACATTAGAAAGACCTTGTTGGGAACA 294
QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
DB 295 -----ACTATGGAAGAAGCTTGTGCTGCGCATGCTCCATGATGGAGCAACGCTAATCAC 348
QY 116 ThrPheSerGIYProLeuSerSerSerIleGIYArgArgProMetLeuIleLeuSerSer 135
DB 349 ACATGCTCTGGCTCATAGCGGATGGCTCGGTCGCGCACCCATATGATTAATCTCATCT 408
QY 136 IleLeuTyrPhePheSerGIYLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu 155
DB 409 GTCCTCATATTTCTTGGGTGGTTTGGTGAATGCTGTGGTCCCAATGTATGTGTGGC 468
QY 136 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
DB 469 TTGGGAGAGCTACTTGATGATTTGGGATTTGGCTTGTGCTGTGATCTGTCTCCGCTCAT 528

QY 176 IleSerGIYIleAlaProSerGIYIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195
DB 529 ATATCTGAACCGCGCGCTGTGAATAAGGGGGTGTGAATACGTTCTTCAGTTCACT 588
QY 196 GlySerGIYIleMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
DB 589 GGCTGTGAGGAAGATTGTTGCTGCTACGTATGTTTGGCATGATGATGACGCCGCCG 648
QY 216 ProAspThrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
DB 649 CTAAGCTGAGAGCTCATGCTGGGCTTGTCTATTCCTTCTCTTGTATTTGCATTG 708
QY 236 ThrIlePheTyrLeuProGluSerProArgTyrPheValSerIleGlyArgMetAlaGlu 255
DB 709 ACCATTTTTTCTTGCCCGAGTCTCCGCGGTGCTGATGACCAAGAGATCTCTGAG 768
QY 256 AlaIysIysValIleGlnIleLeuArgGlyIleAspAspValSerGIYLeuLeuSerLeu 275
DB 769 GCTAAGAGAGTGTCCCAAGATTGCGGAAGAGGATGTGTGAGCGCATGATG 828
QY 276 LeuLeuGluGluValLeuGluValGlyIleAspPheSerIleGluGluTyrIleIleGlyPro 295
DB 829 CTGCTGAAGGTCTCGGAGATTGGGGTGAATCATCTATCGAAGATCATATATTGGCCCT 888
QY 296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspIleGluGlnIleThrLeuTyr 315
DB 889 GCTGACATGTGGTGTGATGTGCATGACATGCAACAGAGAAAGATTAATTCATATAT 948
QY 316 GlyProGluGluGlnIleSerTyrIleAlaArgProSerIleGlyProIleMetLeuGly 335
DB 949 GGAATCCCAAGAGAGCTTCTTGTGTTTCAAAACCTGTCACTGACAGAGTTCAATGGC 1008
QY 336 SerValLeuSerIleAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355
DB 1009 -----CTTGGCTCACCATGAGCAATCATCAACCAAGATATCCCTCATG 1056
QY 356 AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly----- 373
DB 1057 GATCTCTGGTGAACCTGTTTGGTATGATCATATGAAAGCTCCCGAGACAGCAAGA 1116
QY 374 GlySerMetArgSerThrLeuPheProAsnGluSerMetPheSerValThrAspGln 393
DB 1117 GGAAGCATGCGCAACACACTGTTTCCAATTTTGAAGCATGTTGACAGCATGTCGAGCCG 1176
QY 394 HisAlaIysAsnGluGlnIleThrAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413
DB 1177 CATGCTAAATTTGAACAATGGATGGAAGAAAGCTTACAAAGGAAAGCTGAGGACTACATG 1236
QY 414 SerAspGIYAlaGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
DB 1237 TCAGATGCAACCCGTGGGAGCTCCGATGATTAATTTGCACAGTCTTTATATCTCACGCCAA 1296
QY 434 AlaThrGluAlaGluGlyIleAspIleValHisHisGlyHisArgGlySerAlaLeu--- 452
DB 1297 ACAACAAGCCTTGA---AAAGACTTACCTCTCTCTCCATGAGGAGATGCTTGGC 1353
QY 453 SerMetArgArgGlnThrLeuLeuGlyGluGly---GlyAspGlyValSerSerThrAsp 471
DB 1354 AGCATGAGCGCTCAACAGATGCTCATGCAAGGTCAGGTGAGCAAGTGTGATGACAGT 1413
QY 472 IleGlyGlyGlyTyrPheGlnLeuAlaTyrPheTyrSerGluGluGluGlyIleAsnGlyArg 491
DB 1414 ATTGCTGTGCTGCGCAACTGCTGCAATGAAATGACATGATA---GGTGAAGATGCAAAA 1470
QY 492 LysGluGlyGlyPheIleValArgValTyrLeuHisGlnGluGlyValProGlySerArgArg 511
DB 1471 CAACAAGAGAGGCTTAAAGATTTATTAATGATGAGGAGGAGTTTGTGATCTCTGCTGT 1530
QY 512 GlySerIleValSerLeuProGlyGlyIleAspValPheGluGlySerGluPheValHis 531
DB 1531 GGATTCATTTATGATTTCCCGGTGAAGG-----GAATTTTGTCCAG 1572
QY 532 AlaAlaIleValSerGlnSerAlaLeuPheSerIleGlyLeuAlaGluProArgMet 551

[illegible]

PD 28-NOV-2002.

XX 17-JAN-2002; 2002US-0051902.

XX 24-APR-1998; 98US-063044P.

PR 14-APR-1999; 99US-0291922.

XX (ALLEN/) ALLEN S M.

PA (HITZ/) HITZ W D.

PA (KINNEY/) KINNEY A J.

XX (TINGEY/) TINGEY S V.

PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;

XX WPI: 2003-340957/32.

DR P-PSDB; ABU08329.

XX

PT Novel plant sugar transport proteins and nucleic acid encoding the

PT protein useful for producing transgenic plants having altered levels of

PT sugar transport protein -

XX

PS Claim 2; Page 19-20; 56pp; English.

XX

CC The present invention relates to the isolation of Arabidopsis

CC thaliana-like or Beta vulgaris-like sugar transport proteins, and

CC the polynucleotide sequences encoding them. The plant sugar transport

CC proteins of the invention have been isolated from corn, rice, soybean,

CC and wheat. The polypeptides of the invention may be used for altering

CC the level of expression of a sugar transport protein in a host cell,

CC by transforming a host cell with a chimeric construct encoding all,

CC or a portion of the sugar transport protein, in sense or antisense

CC orientation. Particularly, the polypeptides may provide a means to

CC control carbohydrate transport and distribution in plants.

CC ABX93198-ABX93205 represent cDNA sequences encoding Arabidopsis

CC thaliana-like sugar transport proteins.

XX

SQ Sequence 2601 BP; 648 A; 543 C; 644 G; 766 T; 0 other:

Alignment Scores:

Pred. No.: 2,94e-224 Length: 2601

Score: 2674,00 Matches: 522

Percent Similarity: 82,20% Conservative: 92

Best Local Similarity: 69,88% Mismatches: 115

Query Match: 65,62% Indels: 18

DB: 25 Gaps: 8

US-10-051-909-32 (1-800) x ABX93201 (1-2601)

QY 56 MeSeSerGlyValAlValLeuValAlAlValAlAlaSerIleGlyAsnLeuLeuGlnGlyTyr 75

Db 175 ATGAAAGGCGCCCTCTCTGTCCTATGTCGCGCTTCATGTAATTCCTCCAAAGATGG 234

QY 76 ASPAsnAlaThrIleAlaIleAlaValLeuTyrIleValGlyGlnPheGlnLeuGlnAsn 95

Db 235 GATPAAAGCTACCAATGCCCGGGCTAAAGCTTACATTAAGAAAGACCTTGCTTGGGACA 294

QY 96 GluProThrValGlnGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115

Db 295 -----ACATATGGAAGAGCTGTGTGGTGGGCAATGCCGATGGAGCAAGCTAATACC 348

QY 116 Th-PhSeSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135

Db 349 ACATAGCTCTGGTCTTATGACCGATGGCTCGGTCGGCGACCAATGATGATATCTCATCT 408

QY 136 IleLeuTyr-PhSeSerGlyLeuIleMetLeuTyrPseProAsnValTyrValLeuLeu 155

Db 409 GTGCTCATATTTCTGGGCTGTGTGGTGAAGCTGTGTGTCCTCCAAATGTATGTGTGTC 468

QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrIleValProLeuTyr 175

Db 469 TTGGCAGGAGCTACTGATGATGATTTGGGATTTGGCTGTGTGATCTTGTCCGGCTTAT 528

QY 176 IleSerGlnIleAlaProSerGlnIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195

Db	529	ATATGGAACGGCGCGCTGAAATAAGGGGCTGTAATACGCTTCTCAGTTCACT	568
Oy	196	GLYSERGLYMETPHELEUSERTYCYAMEVALPHEGLYMETSERLEUSERPROSER	215
Db	589	GGCTCTGGAGGAATGTTTTGTGTCGTCTATGATGTTTTGGCATGTCATGATGCCGCG	648
Oy	216	PROABTPRARGILEMETLEUGLYVALLEUALLEPROSERLEUPHEPHEPHEGLYLEU	235
Db	649	CCTAGCTGGAGGCTCAATGCTTGAGGGTTCGTCTATTCCTCTCTCTTGAATTTGGCAATG	708
Oy	236	THRILEPHELYLEUPROGLIUSERPROAGTIRPLEUVALSERIYSGIYARPMETALAGLU	255
Db	709	ACCATTTTTTTTCTTGCCCGAGTCTCTCGGTGCTGTCAAGCAAGAGAGATGCTCGAG	768
Oy	256	ALALYLYVALLEUGLINDLYSLEUAARGIYLYASPAAPVALSERIYGLULEUSERLEU	275
Db	769	GCTAAGAAAGTCTCTCAAGATTTGCGCGGAGAGGAGATGTGTCAAGCGGAGATGGCAATG	828
Oy	276	LEULEUGLUGLYLEUGLINDLYVALGILYGLIARPTTHERILEGLINDUTYRIELLEGLYPRO	295
Db	829	CTGTGTGAAGGCTCTGGGATGTGGGGGTGTATCATCTATGAGACAGATCACTAATTTGGCCCT	888
Oy	296	ALATHTGLUALALASAPRASPLEUVALTHIRAPGLYASPLYSGLINDIETHTLEUTYR	315
Db	889	GCTGACGATGTGGCTGATGTGTCATGAACATGCCAACAGAGAAATATAAATTCGATTATAT	948
Oy	316	GLYPROGLUGLUGLYGLINSETRIPRILEALARGPROSELYSGIYPROILEMETLEUGLY	335
Db	949	GGATCCCAAGCAGGCGCTTCTTGTTATATCAAAACGTGTCACTGACAGAGTTCTATTTGGC	100
Oy	336	SERVALLEUSERLEUALASERARGHISGLYSEMERVALASNGINSEVALPROLEUMET	355
Db	1009	-----CTTGCGTCACACCATGAGATCATCAACCAAGACATGCCCCCTCATG	105
Oy	356	ASPPROLIEVALTHLEUPHEGLYSERVALHISGLIASMETPROGLINDLAGLY-----	373
Db	1057	GATCTCTGTGTACACTGTTTGTGTACATTCATGAAAGACCTCCCAAGACAGAGCAAGA	111
Oy	374	GLYSEMERLARGSERTHLEUPHERPROAENPHEGLYSEMERPHESEVALTHIRASGLIN	393
Db	1117	GGAGACATGCGAAGCACTGTGTTCCAAATTTTGGAGAAGATGTTTCAAGCACTGTGAGCCG	117
Oy	394	HISALALYASNGLUGLINTPRASGLIUAENLEUHHISARGYAPRASARGLINDUTYRALA	413
Db	1177	CATGCTAAATATGGAACAATGGGATGAGAAACCTTCAAGGGAGCGTAGAGCACTACATG	123
Oy	414	SERASRGLYALAGLYASPTYRGLIUAERAPLEUHHISERPROLEUSERLARGLIN	433
Db	1237	TCAGATGCAACCCGTTGGGAGACTCCAGATGAAATTTGCAACAGTCTTTTAATCTCACGCCAA	129
Oy	434	ALATHTGLUALAGLUGLYASAPRILEVALHISAGLYHHISARGYUSERALALEU----	452
Db	1287	ACAACAAACCTTGAA-----AAACACTTACCTCTCCCTCCCATGAGCATGATCTTTGGC	135
Oy	453	SERMETARGARGINTHIREULEUGLUGLYY-----GLYASRGLYVALSERSETHRAP	471
Db	1354	AGCATGAGGCGTCACAGTAGTCTCATGCAAGGTCAGGTAGAGCAAGTGTGTAGTACAGGT	141
Oy	472	ILEGLIYGLIYTPRGLINDLEUALATPRPLYSTPRSEGLINDLYLEGLIYGLINDENGLYARG	491
Db	1414	ATGTGTGTGTGCTGGCACTGGCATGGAATGGAATGACATGAATA-----GCTAGAGATGGAATA	147
Oy	492	LYSGIUGLUGLYPHELYEARGVALTYRLEUHHISGLINDGLYVALPROGLIUSERARGATG	511
Db	1471	CACACAGAGGGTTTAAAGATTTATTTTACATGAGAGAGAGACTTTCTGCATCTCGTCCGT	153
Oy	512	GLYSERILEVALSERLEUPROGLIYGLIYASPRVALPHEGLUGLYSERGLIUPHEVALHIS	531
Db	1531	GGATTCATATTTGATCATCTCCCGGTGAAGC-----CAATTTGTTCAG	157
Oy	532	ALALALALALEUVALSERGLINSERLALEUPHESELYSGIYLEUALAGIUPROARGMET	551

Db	1573	CGTGTGCTTGGTGAAGCAACCGGCTCTTTACTTCAAGAGACTTATTGATGACACCA	1632
Qy	552	SerSepala1a1aMetValHisProSerGluValAlaAluysglYSerArgTrpIlyAsp	571
Db	1633	GTTGGAGCTCGCAATGTTCCACCCATCTGAGCAGCTTCAAGGGGCCAAGTTGGAAACT	1692
Qy	572	LeuPhegluPProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln	591
Db	1693	CTTCTTGAAACAGGGGTTTMAACATGCAATGGTGTGTTGAGTGGTAATCAATACTTCAG	1752
Qy	592	GlnPheAlaGlyIleAsnGlyValIleuYrTrpTrpProGlnIleLeuGlnGlnAlaGly	611
Db	1753	CAGTTTTCAGGGAATAAATGGGGTCTTAATTAACACACTTCAATCTTGAAGAAGCCGGT	1812
Qy	612	ValAlaValIleLeuSerIlyPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer	631
Db	1813	GTTGAAGTTCTTCTTGATGATATPAGGATGGCTGACAGTGGCGATCATTCCTTATACGT	1872
Qy	632	SerIleuTrpTrpLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu	651
Db	1873	GCTTTCACAACTTCTTGATGCTTCCCTCGTATPAGCGGACGACGAGACTCATGATGTT	1932
Qy	652	SerGlyArgArgPheLeuLeuLeuGlyTrpIleProIleLeuIleAlaSerLeuValIle	671
Db	1933	TCAGGACAGAAAGCGCTTGCTACTACTACAAATCCCCGGCTGATGTGTCTACTCATTTT	1992
Qy	672	LeuValValSerAsnLeuLeuLeuAspLeuGlyTrpIleuAlaHisAlaLeuLeuSerThrVal	691
Db	1993	TTGGTCATGGAAGCTGTGTAATTTTGGCAATGCGCCCATGCAAGCATCTCAACAGTA	2052
Qy	692	SerValIleValIlyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu	711
Db	2053	TGCGTTGGGTATTATTCTGCTGCTTGTGTATGGGTATGAGCAATTCCAACATCCTT	2112
Qy	712	CysAlaGlnIlePheProThrArgValAlaArgIlyLeuCysIleAlaIleCysAlaPheThr	731
Db	2113	TGCTCAGAGATTTCGCCCATGAGGGTGGTGGCTCTGCATTTGCTATCTTGGCATTATGTG	2172
Qy	732	PheTrpIleGlyAspIleIleValIleValThrTySerLeuProValMetLeuAsnAlaIleGly	751
Db	2173	TTTCGGATGGAAGCATCATCATCATCACTACGCTGCTGTGATGCTCGGCTCTTACGTA	2232
Qy	752	LeuAlaGlyValPheSerIleTyTrAlaValAlaCysLeuIleSerPheValPhe	771
Db	2233	CTGTGGTGTATTCGCAATTACGCAAGTGTGTTGTTTCATCTCGTGAATATTGTGTTT	2292
Qy	772	LeuIlyValProGluTrpIlyrGlyIleMetProLeuGluValIleTrpGluPhePheAlaVal	791
Db	2293	TTGGAAGTTCCAGAAACAAAGGGATGCGCTTBAAGTCATCTCGAATTTCTTCTGTT	2352
Qy	792	GlyAlaIlyrGlnAlaAlaAla 798	
Db	2353	GGAGCAAAAGCGGCTGCTTCT 2373	
RESULT 5			
ABZ13510			
ID	ABZ13510	standard; DNA; 2190 BP.	
AC	ABZ13510;		
XX	21-JAN-2003	(first entry)	
DT			
XX	Arabidopsis thaliana stress regulated gene SEQ ID NO 1315.		
DE			
XX	Arabidopsis thaliana; plant; gene; stress; transgenic; ds.		
XX	Arabidopsis thaliana.		
OS			
XX	WO200216655-A2.		
PN			
XX	28-FEB-2002.		
PD			
XX	24-AUG-2001; 2001WO-US26685.		

Db	1447	GCTATGGACATCCATCCGAA---ACAACTAAAGGCTAATTGGCATGATCTTCATCATG	1503
Qy	575	ProglYValArgargAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAla	594
Db	1504	CCTGGAGTCAGCCGCGCATTAAGTCGTAGAGAGTTGGACTTCAAAATCACTTCAGACAGTTCTCA	1563
Qy	595	GlyIleLeuGlyValLeuTyrTyrThrProGlnIleLeuGlnGlnAlaGlyValAlaVal	614
Db	1564	GGCATCAACGGAGTCTTTACTTACACACCGGAAATCTTGVAGCGCGGCTGTGGGAGTC	1623
Qy	615	IleuSerSerIlePheGlyLeuSerSerAlaSerIleSerIleLeuIleSerSerLeuThr	634
Db	1624	CTCACTATGAAACATGGGGATTAAGTTCTTCCCTCAGCATCTTACTTATAGTGCATTGACA	1683
Qy	635	ThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg	654
Db	1684	ACCTTTTGAGATGTTACCTGCACATAGCTCTTGCAATGAGAGCTCATGAGATCTTTCGTGTGA	1743
Qy	655	ArgPheLeuLeuLeuGlyTyrIleProIleLeuIleAlaSerLeuValIleLeuValVal	674
Db	1744	AGGACCTTGCTTCTCAACCAACATTCCAATCCTGATAGCATCTTAATGGTTTATGATATC	1803
Qy	675	SerAlaLeuIleAspLeuGlyTyrThrLeuAlaAlaLeuLeuSerThrValSerValIle	694
Db	1804	TCAATCTGTTCACATGACATGACATTTGTGACGCGGCTTATCAACCGTAGCGTTGTG	1863
Qy	695	ValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGlu	714
Db	1864	CTCAATCTCTGCTCTTCTGTATGAGGTTTCGGTCTCGTCCAAACATCTCTGTGTTCAGAG	1923
Qy	715	IlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIle	734
Db	1924	ATTTTTCCAATCTCGAGTCGCGGGAATCTGCATCGCATCTCGCGCACTCACCTTCTGGATC	1983
Qy	735	GlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGly	754
Db	1984	TGTGACATTAATCGTCACTTACAGTCTCCCGTGTGCTCAAAATCATTTGAGACTAGCTGGT	2043
Qy	755	ValPheSerIleTyrAlaValAlaCysLeuIleSerPheValPheValPheLeuValVal	774
Db	2044	GTTGTTGGAAATGTACGCAATGCTATGTGCACTTCAATGGGCTTTGTGTTCATTAAGTC	2103
Qy	775	ProGluThrIlyGlyMetProLeuGluValIleThrGluPhePheAlaValAlaAlaLys	794
Db	2104	CCGGAAATTAAGAAGCATGCACCTTAAGTCAATCACAGAGTCTTTCTGTGGAGCTGGA	2163
Qy	795	GlnAlaAlaAla 798	
Db	2164	CAAGCTGAAGCT 2175	
RESULT 6			
ABZ14449			
ID	ABZ14449	standard; DNA, 2205 BP.	
XX	ABZ14449;		
XX	AC		
DT	21-JAN-2003	(first entry)	
XX	DE	Arabidopsis thaliana stress regulated gene SEQ ID NO 2254.	
XX	KW	Arabidopsis thaliana; plant; gene; stress; transgenic; ds.	
XX	OS	Arabidopsis thaliana.	
XX	PN	WO200216655-A2.	
XX	PD	28-FEB-2002.	
XX	PF	24-AUG-2001; 2001WO-US26685.	
XX	PR	24-AUG-2000; 2000US-227866P.	
PR	26-JAN-2001; 2001US-264647P.		

PR	22-JUN-2001; 200US-30011IP.
XX	(SGRI) SCRIPS RES INST.
PA	(SYGN) SYNGENTA PARTICIPATIONS AG.
PI	Harper JF, Kreps J, Wang X, Zhu T;
XX	
DR	WPI; 2002-304127/34.
PT	Identifying a stress condition to which a plant cell has been exposed and producing plants with increased tolerance to these abiotic stresses
XX	-
PS	Claim 144; SEQ ID NO 2254; 577bp + Sequence Listing; English.
XX	
CC	The invention relates to identifying a stress condition to which a plant cell has been exposed, comprising:
CC	(a) contacting nucleic acid representative of expressed polynucleotides in the plant cell with an array or probes representative of the plant cell genome; and
CC	(b) detecting a profile of expressed polynucleotides in the plant cell characteristic of a stress response. The method is useful in the production of transgenic plants, cells and seeds and in producing plants with increased tolerance to abiotic stress. The present sequence is that of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used in methods of the invention.
CC	Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied to Derwent by the European Patent Office.
XX	
SQ	Sequence 2205 BP; 521 A; 524 C; 521 G; 639 T; 0 other;
Alignment Scores:	
Pred. No.:	2, 29e-189 Length: 2205
Score:	2275, 00 Matches: 462
Percent Similarity:	75, 88% Conservative: 111
Best Local Similarity:	61, 19% Mismatches: 146
Query Match:	55, 83% Indels: 36
DB:	24 Gaps: 14
US-10-051-909-32 (1-800) x AB214449 (1-2205)	
OY	56 MetSerGIyAlValleuValAlaleValAlaserIIegIyaenLeuIngInGlyTTP 75
Db	1 ATGAAGGGAGCACTCTCGTGCTCTGCCGCCACAAATCGGCATTTCATAAAGATGG 60
OY	76 AspaAmIAthrIlleAlalAlaValleuTyrlleLysIsGsGuInheGlnleaAn 95
Db	61 GAACAAGGCCAACATTTGCTGGAGACTTAAGTTTATATCAACAAAGCTGAATCTA----- 114
OY	96 GlUpProThr---ValGlUgLyLeuIlleValIserMetSerleuIIegIyalatrIlleVal 114
Db	115 ----CCAACCCTGTTCAGAAGCTTTGCTGCTCATATGATCATTGATCGGTGCAACGTCATC 171
OY	115 ThrThPhseSergLyProLeuSerAspserIIegIylarGarqPrometLeuIIleuSer 134
Db	172 ACAGCTTGCTCAGAGCACCATATCTGATGGCTCGGCAAGCCGCCCATGCTCATTTATCA 231
OY	135 SerIIleuTyRpheserSergLyIleuIIemetLeuTpsrProdenValTYrrValleu 154
Db	232 TCAGTATGATGATTCCTCTGCGGTTTGATATGTTGCGTCTCCCAATGCTATGTTCTG 291
OY	155 LeuLeuAlaRphValAspGIypheGylIlegIleuAlaValThrLeuValProLeu 174
Db	292 TGCTTTCGACGCTTCTTAATAGGTTTGTCGCGGCTCGCGGTACATTGCTTCCCTGTT 351
OY	175 TyrIIseSergLuIIleAlaProSerSergLuIIleargIyleuLeuAenThrLeuProGlnPhe 194
Db	352 TACATTTCGAAAACCGCTCCTCCGAGATACAGAGACAGTAAATATCACTCCCTCACTT 411
OY	195 SerGIyseryGIyImetPheLeuSerTYrcymetValPheGylMetSerleuSerPro 214
Db	412 CTGGGCTCTGGAGAAATGTTTTGTTCATACGTATAGTTTCACTATGTCCTCGAGTAC 471

ID ABX93202 standard; cDNA; 1692 BP.
 XX AC ABX93202;
 XX 29-MAY-2003 (first entry)
 XX DE cDNA encoding soybean sugar transport protein #2.
 XX Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
 XX plant sugar transport protein; carbohydrate transport; soybean;
 XX carbohydrate distribution; plant; gene; ss.
 XX Glycine max.
 XX OS US2002178468-A1.
 XX PN 28-NOV-2002.
 XX PD 17-JAN-2002; 2002US-0051902.
 XX PF 24-APR-1998; 98US-083044P.
 XX PR 14-APR-1999; 99US-0291922.
 XX PA (ALTE/) ALLEN S M.
 XX PA (HITZ/) HITZ W D.
 XX PA (KINNEY/) KINNEY A J.
 XX PA (TING/) TINGEY S V.
 XX PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
 XX DR WPI; 2003-340957/32.
 XX DR P-PSDB; ABU08330.
 XX PT Novel plant sugar transport proteins and nucleic acid encoding the
 XX PT protein useful for producing transgenic plants having altered levels of
 XX PT sugar transport protein -
 XX PS Claim 2; Page 22-23; 56pp; English.
 XX CC The present invention relates to the isolation of Arabidopsis
 XX CC thaliana-like or Beta vulgaris-like sugar transport proteins, and
 XX CC the polynucleotide sequences encoding them. The plant sugar transport
 XX CC proteins of the invention have been isolated from corn, rice, soybean,
 XX CC and wheat. The polypeptides of the invention may be used for altering
 XX CC the level of expression of a sugar transport protein in a host cell,
 XX CC by transforming a host cell with a chimeric construct encoding all,
 XX CC or a portion of the sugar transport protein, in sense or antisense
 XX CC orientation. Particularly, the polypeptides may provide a means to
 XX CC control carbohydrate transport and distribution in plants.
 XX CC ABX93198-ABX93205 represent cDNA sequences encoding Arabidopsis
 XX CC thaliana-like sugar transport proteins.
 XX SQ Sequence 1692 BP; 430 A; 339 C; 400 G; 523 T; 0 other;
 Alignment Scores:
 Pred. No.: 1,636-136 Length: 1692
 Score: 1670.50 Matches: 333
 Percent Similarity: 79.75% Conservative: 57
 Best Local Similarity: 68.10% Mismatches: 84
 Query Match: 40.99% Indels: 15
 DB: 25 Gaps: 8
 US-10-051-909-32 (1-800) x ABX93202 (1-1692)
 QY 308 Asplysglunililethleutyrglproglunglglinsertipilealaargpro 327
 Db 21 GAATAAGATCAATTAAGTTGATGACCAAGAACAGCCAGTCTGGGTTGCTAGACT 80
 QY 328 Serlysglproliemleucliservalleuserialaaserghisglisermet 347
 Db 81 GTTGCTGACCAATATCTGTGGC-----CTTGATCTTAGAAGAAAGGAGCATG 128
 QY 348 Valhenginservalproleumetappprolielvalthrleuphegliservalhisglu 367

Db 129 GCAATCAAGC---AGCTAGTGACCCCTAGTACCCCTTTGGTAGTACATAG 185
 QY 368 Asmetproglinalaglyglisermetargserthrleupheproamphglisermet 387
 Db 186 AAGCTCCCAAGAACAGG-----AGCACCCCTTTTCCACACTTTGGGGGTATG 233
 QY 388 PheSerValThrAspGlnHisAlaLysAsnGluGlnTyrPaspGluGluAsnLeuHisArg 407
 Db 234 TTCAGTTTGGGGGAATACACCAAGCAATAGATTGGATGAGAAAGCTTACCCAGA 293
 QY 408 AspAspGluGluTyrAlaSerAspGlyValaGlyGlyAspTyrGlyAspAsnLeuHisSer 427
 Db 294 GAGGCTGATGATTAATGCTCTGAT-----GCTGGTATCTGATGACAAATTGGCAGCT 347
 QY 428 ProleuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisGlyHis 447
 Db 348 CCATTGATCTCAGCTCAACCAAGAGCTCTGAT---AAGGACATACCTCTCATGCCCAT 404
 QY 448 ArgGlySerAlaLeuSerMetArgArgGlnThrLeu---GlyGluGlyValAspGly 466
 Db 405 AGTAACTTGA---AGCATAGGCAAGTATGCTCTTTACATGGAATTCAGAGAACCC 461
 QY 467 ValSerSerThrAspIleGlyGlyTyrPglInleuAlaTyrPstPserGluLysGlu 486
 Db 462 ACTGGTACTGAGGATTTGGTGTGGTGGCAGCTAGCATGGAATGCTCGAAAGAG 521
 QY 487 GlyLengLysArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGlyVal 506
 Db 522 GGCCAGATGGAAGAAAGAGAGTGGCTTCAAGAAATATATTTAACCACCAAGATGGTGT 581
 QY 507 ProGlySerArgArgGlySerIleValSerLeuProGlyGlyValAspValPheGluGly 526
 Db 582 TCTGATTTAAGCGGGTCTGTGGTTCACTCT---GGCGGATTTACCAACTGAC 638
 QY 527 SerGluPheValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSerGlyLeu 546
 Db 639 AGTGAGTTGTAACAGGCTGCTGCTGGTAGTACAGCTGCCCTTTAATAGAGACCTT 698
 QY 547 AlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaLysGly 566
 Db 699 ATGGCTCAACCGCCAGTGGACCAAGCTATGATATTCCTCTGAAACAAATTGCAAAAGGG 758
 QY 567 SerArgTyrAspLeuPheGluProGlyValArgArgAlaLeuValGlyValGly 586
 Db 759 CCAAGTTGGAATGATCTTTTGAACCTGGGGTGAAGCATGATGATTTGGGGGTGGA 818
 QY 587 IleGlnIleLeuGlnInPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIle 606
 Db 819 ATGCAATTTCTTACAGGTTCTCTGATATAATGGGGTCCCTACTATAGCCCTCAATT 878
 QY 607 LeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAla 626
 Db 879 CTTGAGCAGGAGCGTGTGTTATCTTTCAAGCCTAGGCTTGTGTTACTTCTTCA 938
 QY 627 SerIleLeuIleSerSerIleThrThrLeuLeuMetLeuProCysIleGlyPheAlaMet 646
 Db 939 TCTTTTCTTATTAATGCGGTACCAACTTGTGATAGCTCTCTTATATAGCCATTGCCATG 998
 QY 647 LeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGluGlyThrIleProIleLeuIle 666
 Db 999 AGGCTCATGATATTTCAGCAGAGAGACTTGTGCTCAGTACAAATCCCGTCTAATA 1058
 QY 667 AlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAla 686
 Db 1059 GCACCTCTTCAATATTAAGTCTCTGGAAAGCTTGTGATTTGGGATCCACAGCAAAATGCA 1118
 QY 687 LeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyPro 706
 Db 1119 TCAATCTCAACCAATTAAGTATGCTATTTCTGTTCTTGTGCAAGGGATTTGGACCA 1178
 QY 707 IleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysAla 726

Db 1179 ATTCTAATATACATTGTGACAGATCTTCCCATCTGAGTCTGCTGCAATGCT 1238
Qy 727 ILeCysAlaPheThrPheTrpIleGlyAspIleIleValThrTySerLeuProValMet 746
Db 1239 ATTGTGCTTACCTTTTGGATCTGTATATCATATGACCTACACATCCCATGTATG 1298
Qy 747 LeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyAlaValCysLeuIleSer 766
Db 1299 CTCATTCGTAGAGCTCGCTGCTGTTTGGTATTATGCTGCTGCTTACACCA 1358
Qy 767 PheValPheValPheLeuValProGluThrTySerGlyMetProLeuGluValIleThr 786
Db 1359 TGGGCTTGTCTTTTGAAGTTCCAGAAACCAAGGCGATGCGCATGAGTATCAT 1418
Qy 787 GluPhePheAlaValAlaAlaGln 795
Db 1419 GAGTCTTCTCTGCGAGCAAAACAG 1445

RESULT 9
ABK51968
ID ABK51968 standard; cDNA; 1487 BP.
XX
AC ABK51968;
XX
DT 27-AUG-2002 (first entry)
XX
DE Wheat contig encoding Arabidopsis thaliana-like sugar transport protein.
XX
KW Wheat; Arabidopsis thaliana-like sugar transport protein;
KW carbohydrate transport; grain filling; annual field crop;
KW plant; gene; ss.
XX
OS Triticum aestivum.
XX
FH Key Location/Qualifiers
FH CDS 3..1040
FT /*tag= a
FT /partial
FT /product= "Portion of a wheat Arabidopsis
FT thaliana-like sugar transport protein"
FT /note= "This sequence lacks a start codon"
XX
XX US6383776-B1.
XX
XX PD 07-MAY-2002.
XX
XX PF 14-APR-1999; 99US-0291922.
XX
XX PR 24-APR-1998; 98US-083044P.
XX
XX PA (DUPO) DU PONT DE NEMOURS & CO E. I.
XX
XX P1 Allen SM, Hiltz WD, Kinney AJ, Tingey SV;
XX
XX DR MPI; 2002-451386/48.
XX
XX P-PSDB; AAU97207.
XX
XX PT New nucleic acid encoding plant sugar-transport proteins, useful for
XX preparing transgenic plants with altered carbohydrate distribution
XX
XX PS Claim 3; Column 47-50; 54p; English.
XX
XX CC The present invention relates to the isolation of plant polynucleotide
XX sequences encoding an Arabidopsis thaliana-like sugar transport
XX protein or Beta vulgaris-like sugar transport protein. The
XX polynucleotide sequences are useful for altering the level of sugar
XX transport proteins in plants, i.e. for control of carbohydrate transport
XX and distribution in plant cells, e.g. during grain filling of annual
XX field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
XX carbohydrate flows and sugar transport. The polynucleotide
XX sequences can also be used to isolate cDNA sequences and genes that
XX encode homologues of the new proteins. The present sequence
XX representing a contig assembled from various wheat cDNA clones encodes

CC a portion of an Arabidopsis thaliana-like sugar transport protein.
XX
SQ Sequence 1487 BP; 355 A; 320 C; 373 G; 439 T; 0 other.
Alignment Scores:
Pred. No.: 5,63e-115 Length: 1487
Score: 1424.00 Matches: 281
Percent Similarity: 89.97% Conservative: 24
Best Local Similarity: 82.89% Mismatches: 32
Query Match: 34.94% Indels: 2
DB: 24 Gaps: 2
US-10-051-909-32 (1-800) x ABK51968 (1-1487)
Qy 462 GluGlyValAspGlyValSerSerThrAspIleGlyGlyTyrTrpGlnLeuAlaTrpLys 481
Db 12 GAGGCTGGGAGGCGAGCTAGACAGCATGTGTATGGTGGGGGATGCAATCGCAAGAAA 71
Qy 482 TrpSerGluLysGluGlyValAsnGlyValArgLysGluGlyPheLysArgValTyrLeu 501
Db 72 TGGTCGAGCGACCAAGGCGAGATGCGAAGAGAGAGAGGCTTCAAAAGATCTACTTG 131
Qy 502 HisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyValGly 521
Db 132 CACCAAGAGGGGGTGGCCGACTCAAGAAAGGGGCTCTGTTTTCACCTTCTGATGGGGGT 191
Qy 522 AspValPheGlu---GlySerGluPheValHisAlaAlaLeuValSerGlnSerAla 540
Db 192 GATGCCAGCAAGGGGCGAGGGGTTTATCATGCTGCTGTGGTAGGCACATCGGCT 251
Qy 541 LeuPheSerLysGlyLeuAlaGluProArgMet---SerAspAlaAlaMetValHisPro 559
Db 252 CTTTACTCCAGAGATCTTATGGAAGAGCTATGGGGCGCCGCCCATGATATATCA 311
Qy 560 SerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArg 579
Db 312 TTGAGGCGAGCTCCCAAGGTTTCAATCGAAGATCTGTTTGAACCTGATGGGCGT 371
Qy 580 AlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyVal 599
Db 372 GCATTGTCGTGGGTGGTAATCAGATGCTTCAGCGATGTCGGAATTAATGAGATT 431
Qy 600 LeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPhe 619
Db 432 CTTTACTTACTCTCAAAATTCGTGAGCAAGCTGTGGTCTTCTTCTTCCATCTT 491
Qy 620 GlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeu 639
Db 492 GGCCTCACTTCAAGATCAGATCCATCCATCTTGAATGATGTTCTCACCACTTACTATGCTC 551
Qy 640 ProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyValArgArgPheLeuLeuLeu 659
Db 552 CCAGCATTTGGTGGAGCATGAGACTTATGATATATCTGGAAGAGGTTTCTGCTACTCG 611
Qy 660 GlyThrIleProIleLeuIleAlaSerSerValIleLeuValSerAsnLeuIleAsp 679
Db 612 GGCACATTTCCATCTTATATGACATCTCTAATGTTTGGTGTGGTCAAAATGTTATCAAC 671
Qy 680 LeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCys 699
Db 672 TTGAGTACGAGTGGCCCAAGCTGTGCTCCACAGTTACGTCATGTCCTACTCTGCTGC 731
Qy 700 PheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArg 719
Db 732 TTTGTCATGGGCTTTGGGCCCATGCCCAACATTCATATGCAAGATTTTCCCAACGCA 791
Qy 720 ValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleVal 739
Db 792 GTCCGTGTGTCTGATGCTATTTTGGCCCTCACATTTCTGATTTGACATTAATTTGTT 851
Qy 740 ThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyr 759
Db 852 ACTTACAGCTGCGCTGATGCTGTAATGCTATTTGGTACGGGGGTCTTTGGTATATAT 911

OY 760 AlaValValCysLeuIleSerPheValPheValPheLeuValProGluThrIleGly 779
 DB 912 GCAGCTGTTGTCGATTCCTTGTCTGCTCTACCTTAAGCTCCAGACCAAGGCG 971
 OY 780 MetProLeuGluValIleThrGluPhePheAlaValGlyAlaLeuGlnAlaAla 798
 DB 972 ATGCCCTCGAGTCATCACCGAGTTCTTGGCGTTGGGGCGAAGCAAGCGAGGCC 1028
 RESULT 10
 ABX93204
 ID ABX93204 standard; cDNA; 1487 BP.
 AC ABX93204;
 XX 29-MAY-2003 (first entry)
 DE cDNA encoding wheat sugar transport protein #2.
 XX Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
 KW plant sugar transport protein; carbohydrate transport; soybean;
 KW carbohydrate distribution; plant; gene; 88.
 XX Trilicium aestivum.
 OS US2002178468-A1.
 XX 28-NOV-2002
 PD 17-JAN-2002; 2002US-0051902.
 XX 24-APR-1998; 98US-083044P.
 PR 14-APR-1999; 99US-0291922.
 XX (ALLE/) ALLEN S M.
 PA (HITZ/) HITZ W D.
 PA (KINN/) KINNEY A J.
 PA (TING/) TINGEY S V.
 XX Allen SM, Hitz WD, Kinney AJ, Tingey SV;
 PI WPI; 2003-340957/32.
 DR P-PSDB; AB008332.
 XX Novel plant sugar transport proteins and nucleic acid encoding the
 PT protein useful for producing transgenic plants having altered levels of
 FT sugar transport protein
 XX Claim 2; Page 25-26; 56pp; English.
 XX The present invention relates to the isolation of Arabidopsis
 CC thaliana-like or Beta vulgaris-like sugar transport proteins, and
 CC the polynucleotide sequences encoding them. The plant sugar transport
 CC proteins of the invention have been isolated from corn, rice, soybean,
 CC and wheat. The polypeptides of the invention may be used for altering
 CC the level of expression of a sugar transport protein in a host cell,
 CC by transforming a host cell with a chimeric construct encoding all,
 CC or a portion of the sugar transport protein, in sense or antisense
 CC orientation. Particularly, the polypeptides may provide a means to
 CC control carbohydrate transport and distribution in plants.
 CC ABX93198-ABX93205 represent cDNA sequences encoding Arabidopsis
 CC thaliana-like sugar transport proteins.
 XX Sequence 1487 BP; 355 A; 320 C; 373 G; 439 T; 0 other;
 SO
 Alignment Scores:
 Pred. No.: 5.63e-115 Length: 1487
 Score: 1424.00 Matches: 281
 Percent Similarity: 89.97% Conservative: 24
 Best Local Similarity: 82.89% Mismatches: 32
 Query Match: 34.94% Indels: 2
 DB: 25 Gaps: 2

US-10-051-909-32 (1-800) x ABX93204 (1-1487)
 OY 462 GluGlyGlyAspGlyValSerSerThrAspIleGlyGlyIleValIleValIleVal 481
 DB 12 GAGGCTGGGAGGAGCTAGAGCACTGGTATGCTGGGAGGAGTGGCAATCGCATGAGAA 71
 OY 482 TrpSerGluValGlyGluValGlyValGlyValGlyValGlyValGlyValGlyVal 501
 DB 72 TGTGTGAGGAGCAAGAGGAGATGAGCAAGAGAGAGAGGCTTCAAGAGATCTACTTG 131
 OY 502 HisGlnGluValProGlySerArgGlySerIleValSerLeuProGlyGlyGly 521
 DB 132 CACCAAGAGGAGGAGTGGCCGACTCAAGAGGCTCTGTGTTCATCTTCGAGGAGGCT 191
 OY 522 AspValPheGlu---GlySerGluPheValHisAlaAlaValValSerGlnSerAla 540
 DB 192 GATGCCACGCAAGGAGGAGTGGCTTATACATCTGCTGTGGTAGGCACTGGAGCT 251
 OY 541 LeuPheSerIleGlyLeuAlaGluProArgMet---SerAspAlaAlaMetValHisPro 559
 DB 252 CTTTACTCCAGAGATCTTATGGAAGAGCTATGGCGCGGCTCCAGCATATTCATCA 311
 OY 560 SerGluValAlaAlaValGlySerArgTrpIysAspLeuPheGluProGlyValArgArg 579
 DB 312 TTGAGAGGAGCTCCCAAGGTTCAATCTGAAAGATCTGTTGAACCTGTGTGAGGCGT 371
 OY 580 AlaLeuValGlyValGlyIleGlnIleLeuGlnIlePheAlaGlyIleValGlyVal 599
 DB 372 GCATGTTGCTGCTGGGAGTGGAAATTCAGATGCTTCAAGCTTGGCAATTAAGAGATT 431
 OY 600 LeuTrpTrpThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerIlePhe 619
 DB 432 CTTACTACTACTCTCAATCTTGAGAGCACTGGTGTGCTTCTTCCAACTT 491
 OY 620 GlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeu 639
 DB 492 GGCCTCAGTTGAGATGAGATCCATCTTGATCTGCTTCCACACCTTACTCATGCTC 551
 OY 640 ProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeu 659
 DB 552 CCAAGCATGTGTAGGATGAGATCTTATGATATCTTGAAGAAGATTCTGCTACTG 611
 OY 660 GlyThrIleProIleLeuIleAlaSerLeuValIleLeuValIleValSerAsnLeuLeu 679
 DB 612 GGCACAAATCCCATCTTATGATCTTATGCTTATGTTGGTGTGCTCAATGATATCAAC 671
 OY 680 LeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValIleValIleVal 699
 DB 672 TTGAGTACGAGTGGCCCAAGCTGTCTCTCCACAGTTAGCGTCAATGTTCTTCTGCTGC 731
 OY 700 PheValMetGlyPheGlyProIleProAsnIleLeuValIleValIlePheProThrArg 719
 DB 732 TTGTGCATGGGCTTTGGCCCATCCCAACATCTTATGTGAGAGATTTCCCAACGAGA 791
 OY 720 ValArgGlyLeuValIleAlaIleCysAlaIlePheThrPheTrpIleGlyAspIleIleVal 739
 DB 792 GTCCGTGTGTCTGATCGCATCTTATTTGCCCTCATCTTGTGATTTGAGATATTTGT 851
 OY 740 ThrTrpSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTrp 759
 DB 852 ACCTAACAGCTGCGCTGATCTGAATGCTATGTGTAGCGGAGTCTTGTGTATATAT 911
 OY 760 AlaValValCysLeuIleSerPheValPheValPheLeuValProGluThrIleGly 779
 DB 912 GCAGCTGTTGTCGATTCCTTGTGTGCTCTACCTTAAGGTCCTCCAGACCAAGGCG 971
 OY 780 MetProLeuGluValIleThrGluPhePheAlaValGlyAlaLeuGlnAlaAla 798
 DB 972 ATGCCCTCGAGTCATCACCGAGTTCTTGGCGTTGGGGCGAAGCAAGCGAGGCC 1028
 RESULT 11
 ABX51969
 ID ABX51969 standard; cDNA; 1009 BP.

Db 79 TATGATACCGAGTATTCGGAGCTATTTATTTATGAAAAAGAGCTTAGCCTTA--- 135
 Qy 95 AaenGIuProthraValIGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleVal 114
 Db 136 ---AACGGCTTACAGAAAGCTTCTTGTCTGACGCTCTTGTGGAGGAGATTTGGGC 192
 Qy 115 ThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSer 134
 Db 193 TCAGAGAGCGCGCGAGAGCTGACGCTTTCGAGAAAGAAAAAGCAATTATGACAGCC 252
 Qy 135 SerIleLeuIlePhePheSerGlyLeuIleMetLeuIlePheSerProAsnValIleValLeu 154
 Db 253 GCGCTGCTTTGATAGCGGCTCTTGTGTGCTGACGCGCCCAATACAGAGTCATG 312
 Qy 155 LeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeu 174
 Db 313 GTGCGTTTCCGATCATTTGGAGCTGACGCGAGACATGACGACATGATCCCTT 372
 Qy 175 TyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPhe 194
 Db 373 TATTATCTGAAGTGGCGCCAAACATAAGCGGCGGCTGTCTATCATCATGATAGCTG 432
 Qy 195 SerGlySerGlyGlyMetPheLeuSerIleCysMetValPheGlyMetSerLeuSerPro 214
 Db 433 ATGATTCACGCTCGGATCTCTTCTTAC---ATTGTCATTTACATATTTGCCGAT 486
 Qy 215 SerProAspIlePheGlyIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGly 234
 Db 487 GCCAGAGCGGCGCTGATGCTGTGATGCTGCTGCGCTCATTTCTCTTGTCTT--- 543
 Qy 235 LeuThrIlePheIlePheProGlyLeuSerProArgIlePheValSerGlyArgMetAla 254
 Db 544 ATTGGCATTGTTTATTCGCGAGAGCCCGCTGCTGCTTACACAAATGGCGAAAGAAC 603
 Qy 255 GluAlaIleValIleValLeuGlnIleValLeuArgGlyIleAspAspValSerGlyLeuSer 274
 Db 604 AAACGAGAAATTTCTTGAATAATTTGCGTGCACAAAGATATP----- 648
 Qy 275 LeuLeuLeuGluGlyLeuGluValIleGlyIleAspThrSerIleGluGluIleIleGly 294
 Db 648 ----- 648
 Qy 295 ProAlaThrGluAlaIleAspAspLeuValThrAspGlyAspGlyGluGlnIleThrLeu 314
 Db 648 ----- 648
 Qy 315 TyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerGlyProIleMetLeu 334
 Db 648 ----- 648
 Qy 335 GlySerValLeuSerLeuAlaSerArgHisGlySerMetValaenGlnSerValProLeu 354
 Db 648 ----- 648
 Qy 355 MetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGly 374
 Db 648 ----- 648
 Qy 375 SerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHis 394
 Db 648 ----- 648
 Qy 395 AlaIleAsnGluGlnIlePheAspGluGluAsnLeuHisArgAspAspGluGlyTrpAlaSer 414
 Db 649 -----GATCAGGAA----- 657
 Qy 415 AspGlyAlaGlyGlyAspTrpGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla 434
 Db 657 ----- 657
 Qy 435 ThrGlyAlaGluGlyIleAspIleValHisHisGlyHisArgGlySerAlaLeuSerMet 454
 Db 658 -----ATACAT----- 663

Qy 445 ArgArgGlnThrLeuLeuGlyGluGlyIleAspGlyValSerSerThrAspIleGlyGly 474
 Db 664 -----GATTA----- 669
 Qy 475 GlyTrpGlnLeuAlaTrpIleTrpSerGlyIleGluGluGluGluGluGluGly 494
 Db 670 -----AAAGAACCGGAAAGCAGATGAAGCGGCTGTAAGAGAG--- 708
 Qy 495 GlyPheIleArgValIlePheHisGlnGluGlyValProGlySerArgArgGlySerIle 514
 Db 708 ----- 708
 Qy 515 ValSerLeuProGlyGlyIleAspValPheGluGlySerGluPheValHisAlaAlaAla 534
 Db 708 ----- 708
 Qy 535 LeuValSerGlnSerAlaLeuPheSerIleGlyLeuAlaGluProArgMetSerAspAla 554
 Db 708 ----- 708
 Qy 555 AlaMetValHisProSerGluValAlaAlaIleGlySerArgTrpIleAspLeuPheGly 574
 Db 709 -----CTGTCGAT 717
 Qy 575 ProGlyValaArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAla 594
 Db 718 CCAAGGCGCGCCAGCGCTTATTCAGAGTTTGGAGCTGCTTTTTCACCAATTATTC 777
 Qy 595 GlyIleAsnGlyValLeuIleIleIleIleIleIleIleIleIleIleIleIleIleIle 614
 Db 778 GGAAGCATAGCATCATCTACTATGCGCAAGACCTTACAAAGCTGGA----- 828
 Qy 615 IleLeuSerIlePheGlyLeuSerSerAlaSerAlaSerIleLeu-----Ile 630
 Db 829 -----TTGCGA-----AACTCCGCTTCGATTTAAGCAGCGTCGGAATC 867
 Qy 631 SerSerLeuThrThrIleLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAsp 650
 Db 868 GGCACAGCATGATTTCTCATACATTA-----GTAAGCATTAATAATCATGCAC 915
 Qy 651 LeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuVal 670
 Db 916 AAGATTGGAAGAAAGCGCTTACTGCTATTCGGAAATCGGCGCATGATCAGCTGATC 975
 Qy 671 IleLeuValIleSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThr 690
 Db 976 GTTCTCGCTTAAATAATCTCTTTTCAATAACACTCGCGCTCATGACGACCGCTC 1035
 Qy 691 ValSerValIleValIleIlePheCysCysPheValMetGlyPheGlyProIleProAsnIle 710
 Db 1036 ATTTGTTTAAAGCGGTGTTATGTTGTCTTTCGCGTCACTGGAGACCGGTTGTGGGTG 1095
 Qy 711 LeuCysAlaGluIlePheProThrArgValaArgGlyLeuCysIleAlaIleCysAlaPhe 730
 Db 1096 ATGCTTCTGAATGTTCCCGCTTACCTCAGAGAAATCGGAGACCGGTTTCCACCTTA 1155
 Qy 731 ThrPheTrpIleGlyAspIleIleValThrIleIleIleIleIleIleIleIleIleIle 750
 Db 1156 ATGTTCACAGTTGGAGACACTGATTGTTTCATTAACTTCAATTAATTAATGAAGCGATC 1215
 Qy 751 GlyLeuAlaGlyValaPheSerIleIleIleIleIleIleIleIleIleIleIleIleIle 770
 Db 1216 GGAATCAGTTATTTATTTCTGATTTATGCGCGAGATCGGTATCAAGCGCTTATTTGTC 1275
 Qy 771 PheLeuIleValProGluThrIleGlyIleMetProLeuGluValaIleThrGlu 787
 Db 1276 CGATTTAAGTACAGACAGACAAAGGAAAGAGCTTGAAGAAATTGACAGC 1326
 RESULT 14
 AAKS1964
 ID AAKS1964 standard; cDNA; 870 BP.
 XX

XX	AC		ABK51964;		
XX	DT		(first entry)		
XX	DE	Rice	cDNA clone rls6.pk0003.d5 encoding sugar transport protein.		
XX	KW	Rice; Arabidopsis thaliana-like sugar transport protein;			
XX	KM	carbohydrate transport; grain filling; annual field crop,			
XX	KX	plant; clone rls6.pk0003.d5; gene; ss.			
OS	Oryza sativa.				
PH	Key	Location/Qualifiers			
FT	CDS	8..403			
FT	/tag= a				
FT	/partial				
FT	/product= "Portion of a rice Arabidopsis				
FT	thaliana-like sugar transport protein"				
FT	/note= "This sequence lacks a start codon"				
FN	US6383776-B1				
PD	07-MAY-2002.				
PD	14-APR-1999;	99US-0291922.			
PR	24-APR-1998;	98US-083044P.			
PA	(DUPO) DU PONT DE NEMOURS & CO E I.				
PI	Allen SM, Hitz WD, Kinney AJ, Tingey SV,				
DR	WPI; 2002-451386/48.				
DR	P-PBDB; AAU97203.				
PT	New nucleic acid encoding plant sugar-transport proteins, useful for				
PT	preparing transgenic plants with altered carbohydrate distribution -				
PS	Claim 3; Column 33-36; 5app; English.				
CC	The present invention relates to the isolation of plant polynucleotide				
CC	sequences encoding an Arabidopsis thaliana-like sugar transport				
CC	protein or Beta vulgaris-like sugar transport protein. The				
CC	polynucleotide sequences are useful for altering the level of sugar				
CC	transport proteins in plants, i.e. for control of carbohydrate transport				
CC	and distribution in plant cells, e.g. during grain filling of annual				
CC	field crops (e.g. corn, rice, soybeans, and wheat), and, for studying				
CC	carbohydrate flows and sugar transport. The polynucleotide				
CC	sequences can also be used to isolate cDNA sequences and genes that				
CC	encode homologues of the new proteins. The present sequence encodes				
CC	a portion of a rice Arabidopsis thaliana-like sugar transport protein.				
SQ	Sequence 870 BP; 191 A; 186 C; 206 G; 287 T; 0 other;				
Alignment Scores:					
Fred. No.:	1.03e-39	Length:	870		
Score:	562.00	Matches:	103		
Percent Similarity:	93.75%	Conservative:	17		
Best local Similarity:	80.47%	Mismatches:	8		
Query Match:	13.79%	Indels:	0		
DB:	24	Gaps:	0		
US-10-051-909-32	(1-800) x ABK51964 (1-870)				
Dy	LeuValIleLeuValSerAsnIleAspLeuGlyThrIleuAlaHisAlaLeuLeu	688			
Dd	8 GTTCAACTTAATTTGGTCATAATTCTTGGAAGTAGGCACAAAGGTCAATGCCCTACTG	67			
Dy	SeTrhVasIseValIlleValTyrrPhcYsCySphVaImetClYPheGlYProllePro	708			
Dd	68 TCACAGTAGAGTACTAATCCTACTTCTGTTCTTGTATAGGGTTCCGGGCTATTCCA	127			
Dy	AenIIeLeuCysAlaGluIlephePoTrhArGVaIArGIylEuCYsIIeaIIaICys	728			

Dd		128	AACATTTCTGTGCGAGAGTTTTTCCCGAACCAACCGTTGGCATCTGCATTAGCATCTGT	167
Oy		729	AlApeThrPheTrpIleGlyAspIleIleValThrTySerLeuProValMetLeuAsn	748
Dd		188	GCCCTAACATTCCTGCGATCGGTATTCATTGTGCATFACACCCGCCGTGATCTCAAC	247
Oy		749	AlAlIleGlyLeuAlaGlyValPheSerIleTyrAlaValCysLeuIleSerPheVal	768
Dd		248	GCCATTGGACTCGCTGGAGCTTTTGGAATCTACGACAGTGGTCTGCATCGGCTTCTTG	307
Oy		769	PheValPheLeuLysValProGluThrTyrSglyMetProLeuGluValIleThrGluPhe	788
Dd		308	TTTGTCTTATGTAAGTGTCCGAGACAAGGCGATGCTCTTGAAGTCATCACCGATTTC	367
Oy		789	PheAlaValAlaGlyAlaLysGlnAla	796
Dd		368	TTCCTGTGCGAGCAAAGCAGGCC	391
RESULT 15				
ABX93200	ID	ABX93200	standard; cDNA; 870 BP.	
ABX93200;	AC			
DT	29-MAY-2003	(first entry)		
DE	cDNA encoding rice sugar transport protein #2.			
KM	Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;			
KW	plant sugar transport protein; carbohydrate transport; soybean;			
KM	carbohydrate distribution; plant; gene; ss.			
XX	Oryza sativa.			
OS	US2002178468-A1.			
PN	28-NOV-2002.			
XX	17-JAN-2002; 2002US-0051902.			
PF	24-APR-1998; 98US-083044P.			
PR	14-APR-1999; 99US-0291922.			
XX	(ALLE/) ALLEN S M.			
PA	(HITZ/) HITZ W D.			
PA	(KINN/) KINNEY A J.			
PA	(TING/) TINGEY S V.			
PI	Allen SM, Hitz WD, Kinney AJ, Tingey SV;			
XX	WPI; 2003-340957/32.			
DR	P-PDB; ABU08328.			
PT	Novel plant sugar transport proteins and nucleic acid encoding the			
PT	protein useful for producing transgenic plants having altered levels of			
PS	sugar transport protein -			
PS	Claim 2; Page 18; 56pp; English.			
XX	The present invention relates to the isolation of Arabidopsis			
CC	thaliana-like or Beta vulgaris-like sugar transport proteins, and			
CC	the polynucleotide sequences encoding them. The plant sugar transport			
CC	proteins of the invention have been isolated from corn, rice, soybean,			
CC	and wheat. The polypeptides of the invention may be used for altering			
CC	the level of expression of a sugar transport protein in a host cell,			
CC	by transforming a host cell with a chimeric construct encoding all,			
CC	or a portion of the sugar transport protein, in sense or antisense			
CC	orientation. Particularly, the polypeptides may provide a means to			
CC	control carbohydrate transport and distribution in plants.			
CC	ABX93198-ABX93205 represent cDNA sequences encoding Arabidopsis			
CC	thaliana-like sugar transport proteins.			
XX				

Sequence 870 BP; 191 A; 186 C; 206 G; 287 T; 0 other;

Alignment Scores:

Pred. No.:	1,03e-39	Length:	870
Score:	562.00	Matches:	103
Percent Similarity:	93.75%	Conservative:	17
Best Local Similarity:	80.47%	Mismatches:	8
Query Match:	13.79%	Indels:	0
DB:	25	Gaps:	0

US-10-051-909-32 (1-800) x AEX93200 (1-870)

```
QY 669 LeuValIleLeuValIleSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeu 688
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 8 GTTCACCTGATCTGCTGATATATCTGATGATGGGACCATGTTCAATGCTCAGCTG 67
QY 689 SerThrValSerValIleValIleValIleValIleValIleValIleValIleValIle 708
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 68 TCCACAGTCAGTGCATCTACTCTGCTTCTTGTGTCATGCGGTTCCGCTATTCGA 127
QY 709 AsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCys 728
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 128 AACATTCCTGTGACAGATTTCCGACACCGTTGTCATCTGCAATGACCACTGT 187
QY 729 AlaPheThrPheThrIleGlyAspIleIleValIleValIleValIleValIleValIle 748
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 188 GCCCTAACATTCGTGATCGGTGATATCATTTGACATACACCTCCCGTATGCTCAAC 247
QY 749 AlaIleGlyLeuAlaGlyValAlaPheSerIleIleValIleValIleValIleValIle 768
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 248 GCCATTGACCTGCTGAGTGTGTAATCTACGACGTGCTGCATACCTGCTTCCG 307
QY 769 PheValPheLeuLeuValAlaProGluThrIleGlyMetProLeuGluValIleThrGluPhe 788
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 308 TTGTCTTCATGAGAGTCCGAGACAAAGGCGCATGCTCTTGAAGTCATCACCGAGTTC 367
QY 789 PheAlaValAlaGlyAlaGlyAla 796
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 368 TTCTCTGTGAGCAAGCAGGCC 391
```

Search completed: January 5, 2004, 19:18:26
Job time : 617 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 5, 2004, 18:59:34 ; Search time 7686 Seconds
(without alignments)
4258.093 Million cell updates/sec

Title: US-10-051-909-32

Perfect score: 4075
Sequence: 1 IRSGSWLAQTPTFPDLDR.....PLEVITPEFVAGKQAAAKA 800

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO/US10051909/runat_05012004_094741_22306/app_query.fasta.1.967
-DB=genemb1 -QFMT=fastcap -SUFFIX=rgc -MINMATCH=0.1 -LOOFCU=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=bloms62 -TRANS=human40.cdi -LIST=45
-DOCLALIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=PLO -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10051909 @CGN 1.1.4944 @runat_05012004_094741_22306 -NCPU=6 -ICPU=3
-NO MMAP -LARGEBUFFER -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmb1:.*
1: gb_da:.*
2: gb_htg:.*
3: gb_in:.*
4: gb_om:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_scs:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vl:.*
15: em_da:.*
16: em_fun:.*
17: em_hum:.*
18: em_in:.*
19: em_mu:.*
20: em_om:.*
21: em_or:.*
22: em_ov:.*
23: em_pat:.*
24: em_ph:.*
25: em_pl:.*
26: em_ro:.*
27: em_scs:.*
28: em_un:.*

29: em_vl:.*
30: em_htg_hum:.*
31: em_htg_inv:.*
32: em_htg_other:.*
33: em_htg_mus:.*
34: em_htg_pln:.*
35: em_htg_rnd:.*
36: em_htg_mam:.*
37: em_htg_vrt:.*
38: em_ey:.*
39: em_htgo_hum:.*
40: em_htgo_mus:.*
41: em_htgo_other:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3517	86.3	2824	6	AR208565 Sequence
2	3155	77.4	2665	2	AY165599 Saccharum
3	3052.5	74.9	137140	2	AP005756 Oryza sat
4	2808	68.9	2614	8	HYU534445 Hordeum v
5	2674	65.6	2601	6	AR208568 Sequence
6	2657	65.2	2570	6	AY094465 Arabidops
7	2584	63.4	2190	6	AX506620 Sequence
8	2584	63.4	2190	8	ATH532570 Arabidops
9	2514	61.7	142114	8	AC073166 Oryza sat
10	2514	61.7	300957	8	AE017116 Oryza sat
11	2502	61.4	86710	8	ATP23E12 Arabidops
12	2502	61.4	197859	8	ATCHRI1V83 Arabidops
13	2268	58.1	135406	8	AC136843 Oryza sat
14	2275	55.8	2205	6	AX412656 Sequence
15	2275	55.8	2205	6	AX507559 Sequence
16	2245	55.1	2190	8	ATH532571 Arabidops
17	2217	54.4	2426	8	ATSDGTRPR Arabidops
18	2191.5	53.8	2516	8	AY053446 Hordeum v
19	2144	52.6	143879	8	AC007369 Arabidops
20	2144	52.6	143879	8	AC069251 Arabidops
21	2128	52.2	114918	8	AP004945 Lotus jap
22	2117	52.0	94349	8	ATP26013 Arabidops
23	2093.5	51.4	100693	2	AC121239 Medicago
24	1968	48.3	120538	2	AP004082 Oryza sat
25	1959	48.1	86162	2	AC126786 Medicago
26	1927.5	47.3	122107	2	AC144482 Medicago
27	1670.5	41.0	1692	6	AR208569 Sequence
28	1557.5	38.2	87592	8	AC099739 Oryza sat
29	1557.5	38.2	127098	8	AC144426 Oryza sat
30	1557.5	38.2	154128	8	AP000615 Oryza sat
31	1424	34.9	1487	6	AR208571 Sequence
32	1159.5	28.5	70311	8	AF128457 Oryza sat
33	1140	28.0	77605	8	AF119222 Oryza sat
34	1140	28.0	142852	8	AF161269 Oryza sat
35	1055.5	25.9	145365	8	AC127019 Medicago
36	973	23.9	106246	8	AF528565 Medicago
37	973	23.9	346296	8	AF090447 Zea mays
38	971	23.8	108553	8	AF527809 Sorghum b
39	968.5	23.8	131457	2	AC144718 Zea mays
40	943	23.1	1009	6	AR208572 Sequence
41	587.5	14.4	1518	6	AX535444 Sequence
42	587.5	14.4	8189	1	B8Z92954 B. subtilis
43	587.5	14.4	212610	1	BSUB0019 Bacillus su
44	562	13.8	870	6	AR208567 Sequence
45	539	13.2	1395	6	AX432492 Sequence

RESULT 1

ALIGNMENTS

AR208565
LOCUS AR208565 2824 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 1 from patent US 6383776.
ACCESSION AR208565
VERSION AR208565.1 GI:21509752
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 2824)
Allen, S.M., Hiltz, W.D., Kinney, A.J. and Ringey, S.V.
Plant sugar transport proteins
JOURNAL Patent: US 6383776-A 1 07-MAY-2002;
FEATURES
source location/Qualifiers
1..2824
/organism="unknown"
BASE COUNT 644 a 649 c 745 g 775 t 11 others
ORIGIN
Alignment Scores:
Pred. No.: 2.13e-225 Length: 2824
Score: 3517.00 Matches: 684
Percent Similarity: 95.18% Conservative: 27
Best Local Similarity: 91.57% Mismatches: 34
Query Match: 86.31% Indels: 2
DB: Gaps: 2
US-10-051-909-32 (1-800) x AR208565 (1-2824)
QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
DB 238 ATGGGGGGGGCGCGTATGCTCCCATCGGGCGCTCTATGGCACTGCTGAGGGCTGG 297
QY 76 AspaAnaIaThrIleAlaAlaValLeuValLeuValLeuValLeuValLeuValLeu 95
DB 238 GACATATGCGACATATGCTGAGCGCTCTGATCAATTAAGAAAGAAATTCACCTGAGAGC 357
QY 96 GluProthValGluGlyLeuValLeuValLeuValLeuValLeuValLeuValLeuVal 115
DB 358 GAGCGCTCTGATGAAAGGCTCATCGCGCATGCTTCATTTGGGCGCAACAGTCATCA 417
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyAlaArgProMetLeuLeuLeuSer 135
DB 418 ACATCTCGGGGGCCAGGCGCTGACGCTGTTAGAGGCCATCTGCTGCCCTCGCT 477
QY 136 IleLeuValPhePheSerGlyLeuValLeuValLeuValLeuValLeuValLeuVal 155
DB 478 GTGCTCTACTTCTGATGAGGCTGATGCTTTGCGCGCAATGTCATCTTGTCTC 537
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuVal 175
DB 538 CTGGCAAGGCTCATGATGAGGCTGATGCTTGGGTCACCTGTCCTCTCTAC 597
QY 176 IleSerGluIleAlaProSerGluIleArg--GlyLeuLeuAsnThrLeuProGlnPhe 194
DB 598 ATCTCCGAACATCGACCCGACAGAAATCTTGGGGCTGTTNAAACGTTGCCGAGTTT 657
QY 195 SerGly---SerGlyGlyMetPheLeuSerValYrYrMetValPheGlyMetSerLeuSer 213
DB 658 ATTTGGGTCAGAGGAGGATGTTCTCTCTACTGATGATGTTGGATGTCCTCTAC 717
QY 214 ProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhe 233
DB 718 CCGAAACCTGATGAGGCTCATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 777
QY 234 GlyLeuThrIlePhePheLeuProGlnSerProArgTrpLeuValSerIleGlyArgMet 253
DB 778 GAGCTGACCTGCTTCTTACTGCTGATCAACAGAGTGGCTTGTAGCAAGAGAGATG 837
QY 254 AlaGluAlaValLeuValLeuValLeuValLeuValLeuValLeuValLeuValLeuVal 273
DB 838 GCGGAGCGAAGAGAGATGTCAGAGGCTGCGGAGAGAGAGATGTCAGAGAGAGAG 897

App

QY 274 SerLeuLeuLeuGluGlyLeuGluValGlyValAspThrSerIleGluGluTrpIleIle 293
DB 898 GCTCTTCTAGTGAAGGTTTGGGGTCGGTAAAGATACCTATTNAGATACATCTT 957
QY 294 GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspGlyGluGlnIleThr 313
DB 958 GAGCTGCGCACCGAGGAGCGGATGATCTTAACTGACGGTGTATAGGAACAATACCA 1017
QY 314 LeuValGlyProGluGluGlyValSerTrpIleAlaArgProSerValPheGlyProIleMet 333
DB 1018 CTTATGCGCTTGAAGAGGCGAGTCACTTCTGATGAGGAGGAGGAGGAGGAGGAGGAG 1077
QY 334 LeuGlySerValLeuSerLeuAlaSerArgIleGlySerMetValLeuGlnSerValPro 353
DB 1078 CTGGAGATGATCTTCTTCTGATCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1137
QY 354 LeuMetAspProIleValThrLeuPheGlySerValIleGlyAsnMetProGlnAlaGly 373
DB 1138 CTTATGATCCGATTTGACACTTTTGTGATGTCATGAGAAATATGCTCAAGCTGGA 1197
QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
DB 1198 GGAAGTATGAGAGACACATTTGTTCCAACTTTGGAAGTATGTTCAAGTGCACAGATCAG 1257
QY 394 HisAlaValAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTrpAla 413
DB 1258 CATGCCAAATATGACAGTGGAGTGAAGATCTTCAATAGGAGTGAAGAGAGTGAAGCA 1317
QY 414 SerAspGlyAlaGlyGlyAspTrpGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
DB 1318 TCTATGATGTCAGAGAGGATGATATAGAGACAAATCTCCATGCCCCATGCTGCTCCAGGAG 1377
QY 434 AlaThrGlyAlaGluGlyValAspIleValHisGlyValAspGlySerAlaLeuSer 453
DB 1378 GCAACAGGTGCGAAGGAGAGACATTTGTCACCATGTCACCTGGAAGTCTTTGAGC 1437
QY 454 MetArgArgGlnThrLeuLeuGlyGlyValAspGlyValSerSerThrAspIleGly 473
DB 1438 ATGAGAGGCAAGCCTTTTGGAGAGGTGAGATGCTGTGACAGCATGATATGCT 1497
QY 474 GlyGlyTrpGluLeuAlaTrpValSerSerGlyValGlyGluGluGlyValGlyValGly 493
DB 1498 GGGGATGCGCAGCTTGTGTAATGTCAGAGAAAGAGTGAATGTAAGAAAGAA 1557
QY 494 GlyGlyPheValArgValTrpLeuHisGluGluGlyValProGlySerArgArgGlySer 513
DB 1558 GGTGATTCAAAGAGTCTTCTGACCAAGAGGAGTCTGCTCAAGAGGAGGCTCA 1617
QY 514 IleValSerLeuProGlyGlyValAspValPheGluGlySerGluPheValHisAlaAla 533
DB 1618 ATGTTTCACTTCCCGGAGGAGGATGCTTGAAGGATGATGATGATGATGATGATGAT 1677
QY 534 AlaLeuValSerGlnSerAlaLeuPheSerIleGlyLeuAlaGluProArgMetSerAsp 553
DB 1678 GCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1737
QY 554 AlaAlaMetValHisProSerGluValAlaAlaValGlySerArgTrpValAspLeuPhe 573
DB 1738 GCTGCAATGTTTCAACCATGATGAGTATGCTCCAAAGTTCACCTTGAAGAAATTTGTT 1797
QY 574 GluProGlyValArgAspAlaLeuValGlyValGlyIleGlnIleLeuGlnPhe 593
DB 1798 GAACTGAGAGAGGCGGCTGTTAGTGTGATGATGATGATGATGATGATGATGATGAT 1857
QY 594 AlaGlyIleAsnGlyValLeuValTrpThrProGlnIleLeuGluGlnAlaGlyValAla 613
DB 1858 GCTGATATTAACGATGTTTCTGATATTAACCAAAATTTGTAGCAAGGAGGATGCA 1917
QY 614 ValIleLeuSerValPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
DB 1918 GTTATTTCTTTCCAAATTTGCTCAGCTCGGCAATCACATCTTGAATCAGTTCTCTC 1977
QY 634 ThrThrLeuLeuMetLeuProGlyIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653


```

Db 1978 ACTACCTACTAATGCTTCCTTGCACTTGCGCTTGCCATGCTGCTATGATCTTCCGCA 2037
Qy 654 ArgAghPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
Db 2038 AGAAGGTTTGTGCTGCTAGGCAACAATCTGATGATGACATCTCTGTTATCTGCTGTT 2097
Qy 674 ValSerAenLeuIleAspLeuGlyThrIleuAlaIleAlaLeuSerThrValSerVal 693
Db 2098 GGTGTCATCTATGATTTGGGTACACAGCCCATGCTTGCTGCTCCACATGCTGTT 2157
Qy 694 IleValIyrPheCyCyPhePheValMetGlyPheGlyProIleProAlaIleLeuCyAla 713
Db 2158 ATCGCTACTTCTGCTGCTGCTGCTTATGGGATTTGGTCCATCCCAACATTTATGTCGA 2217
Qy 714 GluIlePheProThrArgValArgIleuCyAlaIleAlaIleCyAlaPhePheThrP 733
Db 2218 GAGATCTTCCACACAGGAGTGTGCTGCTGATTTGCCATTTGCTTACATTTCTGG 2277
Qy 734 IleGlyAspIleIleValThrIyrSerLeuProValMetLeuAlaIleGlyLeuAla 753
Db 2278 ATCGAGATATCATGCTACCTACACAGCTTCTGATGCTGATGCTATGCTGACTGGCG 2337
Qy 754 GlyValPheSerIleIyrAlaValAlaValCysLeuIleSerPheValPheValPheLeuIyr 773
Db 2338 GGTGTTTTCAGCATATGACAGTCTGATGCTGATTTCTTGTGCTGCTTCTCTAG 2397
Qy 774 ValProGluThrIyrGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
Db 2398 GTCCCTTGAGACAAAGGAGTGCCTTGAGATTATACGAATCTTTGACAGTTGATGCG 2457
Qy 794 LysGlnAlaAlaAlaAlaAla 800
Db 2458 AAGCAAGCGCTGCAAAAGCC 2478

RESULT 2
LOCUS AY165599 2665 bp mRNA linear PLN 07-JUN-2003
DEFINITION Saccharum hybrid cultivar putative sugar transporter type 2a mRNA,
complete cds.
ACCESSION AY165599
VERSION AY165599.1 GI:31505503
KEYWORDS
SOURCE
ORGANISM Saccharum hybrid cultivar (sugarcane)
SACCHARUM hybrid cultivar
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum.
1 (bases 1 to 2665)
Caau,R.E., Grof,C.P.L., Rae,A.L., McIntyre,C.L., Dimmock,C.M. and
Mannere,J.M.
Identification of a novel sugar transporter homologue strongly
expressed in maturing stem vascular tissues of sugarcane by
expressed sequence tag and microarray analysis
Plant Mol. Biol. 52, 371-386 (2003)
2 (bases 1 to 2665)
Caau,R.E., Grof,C.P.L., Rae,A.L., McIntyre,C.L., Dimmock,C.M. and
Mannere,J.M.
Direct Submission
Submitted (17-Oct-2002) Division of Plant Industry, Commonwealth
Scientific and Industrial Organisation, 120 Meiers Rd,
Indooroopilly, QLD 4068, Australia
location/Qualifiers
1. 2665
/organism="Saccharum hybrid cultivar"
/mol_type="mRNA"
/db_xref="taxon:128810"
97. 2334
/codon_start=1
/product="putative sugar transporter type 2a"
/protein_id="AA037640.1"
/db_xref="GI:31505504"
/translation="MSGAAVAIVASIGNLQGMNATISALLYIKKEFKLESPTV

```

```

EGLIIVANSLIGATITITSGFVSDWIGRLPMLILSVLYFVSSILIMLSPNVVLLLA
RLINGFVGIALVTLVPLVYISLTAPEIRLGLINTLPQSGSGAMFLSYCMFGMLPS
PDMRIMGVIALPLPFFFGILTYLPLSPRVLVKGMAEKAKVLORLGRDVTGEM
ALIVEGLVEGGDTPIEEVYIGPAIEPDHVAODRRILLFGEEBOSTVARSKSS
SMLOSUSISGRHSKMNQNPILDTPIVTLFGSVHENSILAGSKMRIVSKPSMIS
ITDQHPRTQWDEENLRHNEVYATDAASVYENNVSPILSRQNTAEGKGIAMHG
RGSALGLRRSLDEGEAAVSTGIGWQOLAMWSREDEGDKKESFKRIYAHQEG
VAGSRGSVSLPGGDASBEKFIHAALVYQSALYPRDITTEQRMAGPATMHPSEA
AKVSRMDLPEPGVRAVLVIGIOLIQFAGINGVLYTPQIMEAGVAVLISNLGL
SSASASLISVTRVLMPLPSIGLAMRMDVSGRFFLLSTPVLVLSIIVLVSNVIE
LGIIVHVLSTISVTLCCPKMGFGIPNLVCEFPTRRGICIAICALIFVGD1
IVTVSLPVMNALIGBVFYIVAVACAIFVYLVKVBETKMPLLEVITTEPFAVAKQ
AVAKA
97. 156
/note="putative; transmembrane-region site"
115. 774
/note="Pfam00083; Region: sugar transporter"
238. 297
/note="putative; transmembrane-region site"
313. 372
/note="putative; transmembrane-region site"
406. 462
/note="putative; transmembrane-region site"
511. 570
/note="putative; transmembrane-region site"
589. 654
/note="putative; transmembrane-region site"
658. 1794
/note="Region: large cytoplasmic loop"
1624. 2298
/note="Pfam00083; Region: sugar transporter"
1672. 1731
/note="putative; transmembrane-region site"
1813. 1869
/note="putative; transmembrane-region site"
1897. 1953
/note="putative; transmembrane-region site"
1969. 2022
/note="putative; transmembrane-region site"
2089. 2151
/note="putative; transmembrane-region site"
2191. 2247
/note="putative; transmembrane-region site"
misc_feature 650 a 561 c 691 g 763 t
ORIGIN
BASE COUNT 650 a 561 c 691 g 763 t
ALIGNMENT Scores:
Pred. No.: 2.93e-201 Length: 2665
Score: 3155.00 Matches: 601
Percent Similarity: 89.56% Conservative: 68
Best Local Similarity: 80.46% Mismatches: 78
Query Match: 77.42% Indels: 0
DB: 8 Gaps: 0
US-10-051-909-32 (1-800) x AY165599 (1-2665)
Qy 54 AspIyMetSerGlyAlaValAlaValAlaValAlaSerIleGlyAsnLeuLeuGln 73
Db 91 GAAAACATGTCGGAGCGCTCTTGTGCAATAGACTTGATGCGGAATCTGCTTAG 150
Qy 74 GlyTPAsPAsnAlaThrIleAlaAlaValLeuIyrIleIyrGlyGlnLeu 93
Db 151 GGGTGGCAATGCCACATCTCAGCTGCTATTTATATTAAGAGAAATTCAAACTA 210
Qy 94 GluAspIyrProThrValIleGlyIleuIleValSerMetSerLeuIleGlyAlaThrIle 113
Db 211 GAAAGTGAGCTACTGTCGAGGAGGCTGATTTGGCGCATGTCATTAATTGAGCAACTATC 270
Qy 114 ValThrThrPheSerGlyProLeuSerAspSerIleGlyAlaGAGPProMetLeuIleu 133
Db 271 ATTACAAATCTCTGACGACATACAGATGAGATGCGGCTCCCTATGCTCATCTCTC 330
Qy 134 SerSerIleLeuIyrPhePheSerGlyLeuIleMetLeuTProSerProAsnValIyrVal 153

```

Db 331 TCTTAGTCTCTACTTTTGTAGCAGCTCATTTAGCTATGCTCCCGCATCTATGTA 390
 Qy 154 LeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValPro 173
 Db 391 CTACAGCTGTGACGCTTATTAATGATTCGGGTGTGGCTGTGCTGTGCTGTGCTGT 450
 Qy 174 LeuYrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAlaThrLeuProGln 193
 Db 451 CTGTCATTTCAGAAACAGCTCTCCAGAGATTAAGGTTTGTCTGAAATACATGCGACAG 510
 Qy 194 PheSerGlySerGlyGlyMetPheLeuSerYrYrCysMetValPheGlyMetSerLeuSer 213
 Db 511 TTCAGTGGGTACGGGGGAGATGTTTTGTCTGATGCTGATGCTGTGGATGCTGCTG 570
 Qy 214 ProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhe 233
 Db 571 CCATCACCTGATTTGAGAGATTAATGCTTGGGGTCTGCGTTACCTTATGTTTCTTT 630
 Qy 234 GlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerIleGlyArgMet 253
 Db 631 GGTTCACAAATTTTATCTGCTGATACCAAGATGCTTGTGACGAAAGGGGAGATG 690
 Qy 254 AlaGluAlaIleValLeuGlnIleLeuAlaArgGlyValAspAspValSerGlyGluLeu 273
 Db 691 GCAGAGGCAAGAGAGGTGTGCAAGATTAACGAGGAGATGACCTCACAGGCGAATG 750
 Qy 274 SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspTrpSerIleGluIleLeuIle 293
 Db 751 GCTCTCTCTTGTGAAGGGTGTGAGGTGCGAGGGGATACCTTCAATGAAGATACATC 810
 Qy 294 GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspIleGluIleThr 313
 Db 811 GGTCTGTCTATGAACACAGCTGATGACACGCTGATGATGATGATGATGATGATG 870
 Qy 314 LeuYrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerIleGlyProIleMet 333
 Db 871 CTATTTGGGCTGAAAGAGGCGCAATCATGATGATGATGATGATGATGATGATG 930
 Qy 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
 Db 931 CTGGAAGGTGCTTCTATCGGATCTGTCATGAGCATGATGATGATGATGATGATG 990
 Qy 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnIleGly 373
 Db 991 CTTAGGATCCATATGTCAAGCTTTTGGAGTGTCCATGAGAAACATCTCTCTGCTGA 1050
 Qy 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
 Db 1051 GGAAGTATGAGAGATAGTGCCTTAAATTTGACAGATGATGATGATGATGATGATG 1110
 Qy 394 HisAlaIleValAsnGluIleTrpAspGluGluAsnLeuHisArgAspAspGluIleVal 413
 Db 1111 CAGCCCAAGACCGAGAGAGTGGATGAGGAACTTCATAGGATTAATAGAGATGATG 1170
 Qy 414 SerAspGlyAlaGlyIleAspTrpGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
 Db 1171 ACTGATGGGTCTGCAAGTACTATGAAACAAATGTACACAGCCCTGCTGCTCCGACAG 1230
 Qy 434 AlaThrGlyAlaGluGlyValAspIleValHisGluIleArgGlySerIleAlaLeuSer 453
 Db 1231 AATACAGTCCGAGAGAGGAGGAGCATGACATCAGGCGACCGTGAAGTCTTTGGGT 1290
 Qy 454 MetArgArgGlnThrLeuLeuGlyGluGlyAspGlyValSerSerThrAspIleGly 473
 Db 1291 TTGAAGAAGAAGAGCTGTCCGATAGAGGTGTGAGGCAAGCAACAGCATGCTGGCTGT 1350
 Qy 474 GlyGlyTrpGlnLeuAlaTrpIleTrpSerGlyIleValGluGlyValArgIleGly 493
 Db 1351 GCGGATGGCAGCTGCGCTGGAATGTGTGAGCGGAGGAGTGTGAGGACGCTGAAGAGGA 1410
 Qy 494 GlyGlyPheIleValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer 513

Db 1411 GGAAGTTCAAAAGATCTACTTACACCAAGAGGAGTGTGCTGCTCAGAAGGGGATCT 1470
 Qy 514 IleValSerLeuProGlyGlyIleAspValPheGluGlySerGluPheValHisAlaIle 533
 Db 1471 GTTGCTCACTTCTGCGGAGGATGATGCTCTGAGAGGTGATGATGATGATGATGATG 1530
 Qy 534 AlaLeuValSerGlnSerAlaLeuPheSerIleGlyLeuAlaGluProArgMetSerAsp 553
 Db 1531 GCTGTGTGACAGCTGCTCTTTTACCCGAGGAGTATTACAGAACAGCGATGGCGGT 1590
 Qy 554 AlaAlaMetValHisProSerGluValAlaAlaIleValGlySerArgTrpIleAspLeuPhe 573
 Db 1591 CCAGTCAATGACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1650
 Qy 574 GluProGlyValAlaArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnIlePhe 593
 Db 1651 GAACCTGTGTGAGACGCGCTGTTGGTGTGATGATGATGATGATGATGATGATGATG 1710
 Qy 594 AlaGlyIleAsnGlyValLeuTyrTrpProGlnIleLeuGluGluAlaGlyValAla 613
 Db 1711 GCGGGAATTAATGGGGTCTCTACTACCTCCGAAATTAATGAGCAAGCTGTGTGCA 1770
 Qy 614 ValIleLeuSerIlePheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
 Db 1771 GTTCTATTTCATCTTGTGCTCAGTTCAGTTCAGTTCAGTTCATCTTATGATGTTCTG 1830
 Qy 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
 Db 1831 ACTGCCCTACTTATGCTCTCCATGATGATGATGATGATGATGATGATGATGATG 1890
 Qy 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
 Db 1891 AGAAGTTCGTGCTGCTAAGCAATTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 1950
 Qy 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
 Db 1951 GTGCTAATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2010
 Qy 694 IleValIlePheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
 Db 2011 ATCACTACTATATCTGCTCAAAATGGGCTTTGGCCCATTCACATTCATATGTC 2070
 Qy 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733
 Db 2071 GAGTTTTCACAGGAGGTTGCGATGTCGATGTCGATGTCGATGTCGATGATGATG 2130
 Qy 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
 Db 2131 GTTGAAGCATTTATGTCATACATGATGATGATGATGATGATGATGATGATGATG 2190
 Qy 754 GlyValPheSerIleTyrAlaValAlaCysLeuIleSerPheValPheValPheLeuIle 773
 Db 2191 GCGGTTTGGCATATACGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2250
 Qy 774 ValProGluThrIleGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
 Db 2251 GTTCCGAGGAAAGAGGAGATCCCTTGAAGTATCATCACTGATGTTGTCAGTTGTGCA 2310
 Qy 794 LysGlnAlaAlaAlaAlaAla 800
 Db 2311 AAGCAAGCGTTGCCAAGGCC 2331
 RESULT 3
 AP005756/c 137140 bp DNA linear HTG 19-SEP-2002
 LOCUS Oryza sativa (japonica cultivar-group) chromosome 2 clone
 DEFINITION OSINBD0035N08, *** SEQUENCING IN PROGRESS ***
 ACCESSION AP005756
 VERSION AP005756.1 GI:23200615
 WORDS HTG, HTGS, PHASE2.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE	AUTHORS	TITLE
1	Sasaki, T., Matsumoto, T. and Katayose, Y.	<i>Oryza sativa</i> nipponbare (GA3) genomic DNA, chromosome 2, BAC

Published Only in Database (2002)

AUTHORS
Sasaki, T., Matsumoto, T. and Katayose, Y.

JOURNAL
Submitted (18-SEP-2002) Takuji Sasaki, National Institute of
Advanced Industrial Science and Technology, Tsukuba, Ibaraki, Japan

2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: tinaakr@niss.affrc.go.jp, tmi@bttv.//com.affrc.go.jp/

COMMENT: It currently consists of 1 cottage. Gars between the cottages

are represented as runs of N. The order of the pieces is to be correct as given, however the sizes of the gaps bet

are based on estimates that have provided by the submitters will be replaced by the finished sequence as

* NOTE: This is a 'working draft' sequence.
available and the accession number will be preserved.

* by the finished sequence as soon as it is available and

Location/Qualifiers

/organism=

/cultivar="Nipponbare"

```
/chromosome="2"
```

BASE COUNT	40995 a	27904 c	27671 g	40
------------	---------	---------	---------	----

9
8
7
6
5
4
3
2
1

0
1
2
3
4
5
6
7
8
9

Pred. No.:	6.23e-192	Length
Score:	3053 50	Match

Percent Similarity:	50.81%	CONC
Best Local Similarity:	47.56%	Mism

Query Match:	74.91%	Index
DB:	2	Games

US-10-051-909-32 (1-800) X AP005756 (1-137140)

3 SerGlySerTyrLeuAlaValGlnThrProPh

Db 88617 GCTGCCCTCCTGGCTTCCATTGAGCCGCTT

23 LeuLeuProSer-----Qy

Db 88557 CCGCTTCCCTTCCCTCGGTTCTTGGCCGTT

QY 27 -----

Db 88497 ATGGAGTTCGCAAGATTTTGGAGTCCTGT

QY 35 LeuProAlaSer-----

Db 88437 GGTGTGTTCAATTCCTTGCTGCTGATTCCTTGC

49 ASPASPILEUGLUASPLYSMETSERGLYA

DB 88383 -----AAGATGTCGGGCG

03 GLYASHEULEUINGIYIIPABPABILALD

[illegible]

Oy	270	-----	270
Db	87141	TTGACAGAACTTAACATATGACAGAAATAGTGTATTTTGAAAAAAGGCTAAAGAAA	87082
Oy	270	-----	270
Db	87081	TCAAATCTATTAAGCGCCTTCTATGSGTGTCTTTTGAAAAAGCTTATGAATCTTCAG	87022
Oy	270	-----	270
Db	87021	TMAAATTTACTAGTAGACATTAACACTACATAATTCATTTTCATAGAGGCAAAATG	86962
Oy	270	-----	270
Db	86961	AATGTACATCTTATCTCACACAGTAGAGCCAAATAGAGTGTCCAACTTTTGCTGA	86902
Oy	270	-----	270
Db	86901	GTCCTAATCTGAGAAAAAATATATGAGTATCTAGATAGACATCGCCCTCTTTT	86842
Oy	270	-----	270
Db	86841	TTTGTGTAAGGTAACCTAGAAATTTGGGGGGTGACCTGTCTATATATGTGATG	86782
Oy	270	-----	270
Db	86781	GGCCGACCCACGGGTCTTATCTATCTAGTTAACTAAATGAATAATTTTGCTTTG	86722
Oy	270	-----	270
Db	86721	AGATTTTAAAGTGAATTAAGTCCATCTAGTTGACTAGAGTGGTCAAGAGCCCCAC	86662
Oy	270	-----	270
Db	86661	AGCGGCATGCTTAACAACAGTAGTGTTCAGAAACGTATGATGATTTATCCTTAAGT	86602
Oy	270	-----	270
Db	86601	TCATGCTTCTCAAACTAATTAACATAATACATGTCACATCAAAATAGCATATGCT	86542
Oy	270	-----	270
Db	86541	GTATGCAAGTGTCTCTATTTCTCTCTCTACACAGACTAAATGACAGTATTCGGTT	86482
Oy	270	-----	270
Db	86481	GGCTGAGACATTAATATGTCTGTATTTTAAATTATCGAACTTCGTGAAGCTGGTG	86422
Oy	270	-----	270
Db	86421	TCGCCAGCTGACTATGCTATCTTCTATTAAGCTGGGCCCTTAATATTCGTATTTG	86362
Oy	270	-----	270
Db	86361	TTTCTGGCAACCAATTCATATTTATTTATTCCTGCTTATTCATATGTTACTACTATTTANG	86302
Oy	270	-----	270
Db	86301	CCAAATGTTTATACCAATATGTCCAAATTAATATCATGTATCTACGTGAAAAAATTAAGTTC	86242
Oy	271	-----	281
Db	86241	ATAATTCATTCAAATTAATCAATTGCGAGAGAAATGCGCTCTCTGTTGAAGTTGGA	86182
Oy	281	-----	301
Db	86181	GGTTGGAGCTGACCTTCATTTGAAGAGTACATCATCGACCGGCTATAGAGCCAGCTGA	86122
Oy	301	-----	321
Db	86121	TGAGCATCTTTATATGCGCATAGACCAAAATACCTGTACGGGCTGAAAGAGGCCA	86062
Oy	321	-----	341
Db	321	-----	341

D	86061	ATCATGGATTGGCTGACCTTCCAAAGGACCCAGCAATCTTGGAGTGTCTTCTTAC	86002
Q	341	aseSerGhiGlySerMetValIasnGlnSerValProLeuMetAspProIleValThre		361
D	86001	ATCTGTCATGGAGCAGCTGATGTGAACCAAGCCGTCACCTTAATGATTCATGTCACGCT		85942
Q	361	uPhGlySerValHisGluAsnMetProGlnIaGlyGlySerMetArgSerThrLeuPh		381
D	85941	TTTGGCAGAGTGTCCATAGAACATGCTCATGTGGAGAGATGCGGAGTACATGTT		85882
Q	381	eProAsnPhGlySerMetPheSerValThrAspGlnHisAlaIysAsnGluGlnTrpAs		401
D	85881	TCCCAACTTTGGTATGATGATGTTTATGTGTGACAGATCAAGCACCCCAAGCTTGAATGGGA	85822
Q	401	pGluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyIAspTyr		421
D	85821	TGAAGAACACTTCATCATGGGATGATGAGAAATACCGCTGTGATGGTGTGAGAGTGACTA		85762
Q	421	rgLAspAsnLeuHisSerProLeuMetSerArgGlnIaIaTrnGlyAlaGluGlyIAsAs		441
D	85761	TGAAGAACAAATGTCCACAGCCCGTGTGCTCCACAGACCCAGCCAGCGCAAGAGGAAGGA		85702
Q	441	pIleValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuGlu		461
D	85701	CATTGCACACCAACCTCATCTGTGGAAGTGTCTGAGTATGAGAGAGAAAGCCTTGGGA	85642
Q	461	yGluGlyIAspGlyValSerSerThrAspIleGlyGlyIYrTrpGlnLeuAlaTrpIly		481
D	85641	AGAGGGTGGCAGGCGGTGAGCAGCACTGGCATTTGGTGGGGATGCACTTGACATGGAA		85582
Q	481	eTrpSerGlnLysGluGlyGluAsnGlyArgIlyGlyIYrPheLysArgValTyrIle		501
D	85581	ATGTCTCAGACCGAAGACCGAGATGGTAAAGAAAGAGATTTAAAGAAGATCTACTT	85522
Q	501	uHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGly		521
D	85521	GCACCAAGAGAGAGTGGCCAGATTCMAAGAAAGGAGCTCATATTTCCTTCTGTGGAGG	85462
Q	521	yAspValPheGluGlySerGluPheValHisAlaAlaIleValSerGlnSerAlaIle		541
D	85461	GGAAGCTCCGGAAGGAGCGCAATTCATACATGCTGCGCTTGGTATGACCAACAGCACT		85402
Q	541	uPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerIle		561
D	85401	TTACTTCCAAAGATATCTTGAAACGGGTATGTCCGGTCCAGCCCATATTCAATCCATCACA		85342
Q	561	uValAlaIleLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaIle		581
D	85341	GGCAGCTGCCAAAGGTTCAAAGCTGAGAAAGTTTGTTGAACCTGAGTACGGGTGCGTT		85282
Q	581	uLeuValGlyValGlyIleGlnIleLeuGln		591
D	85281	GTTAGTGGGTGTGGAAATTCATAATCCTTCAACAGGTAAAGATGATACCAATTCATAATT		85222
Q	591	-----		591
D	85221	CTTTTGTGACTGTCCAAACATGACGGTGTTTAAGATTGTGTACTATTAAAGAAATCACT		85162
Q	592	-----GlnPheAlaGlyIleAsnGlyValLeuTyrTyrT		603
D	85161	AATTCTTTATATGTTCTTCTCTTGTGACAGTTGTCTGAAATTAATATGGGGTTCTCTATTACA		85102
Q	603	hrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerS		623
D	85101	CTTCCACAAATATCTTGAGCAACGGGGGTGGCAAGTCTCTCTTCCATCTTGGCTCAAGT		85042
Q	623	eAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleG		643
D	85041	CAGCATCAGGCTTCATCTTATCATGTTCTGTGACCAACCTTACGATGCTTCCAGACATTG		84982
Q	643	lyPheAlaMetLeuLeuMetAspLeuSerGlyArg-----		654

Db		84961	GTTTAGGCATGAGCTTATGGACATCTCTGGAAAGAAGTACTAATTCTTGTCCTTGTTT	84962
Oy		654	-----	654
Db		84921	GCATCATCTTAGTACTACCAAAATATACCAAAGAATGACCACCAAGTAGATTAAATT	84862
Oy		655	-----	
Db		84861	ACATCATACATTGTTCCTTTTAGGTTTCCTTCCTCGCCACCAATTCAGTCTTATAG	84802
Oy		667	IAsErLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaL	667
Db		84801	CATCTCTAGTTGCTTGTGGTTGTGTGCCATGTAATCAGCTCGGGTACAGTGGCCCCAGCGG	84742
Oy		687	euleuSerThrValSerValIleValIlyrPheCySCysPheValMetGlyPheGlyProI	707
Db		84741	CACCTCCACACATTGTGTGCGGAAGATCTTCCCACATAGGCGCCGCGGATCTGCATGGCA	84682
Oy		707	IeProAsnIleLeuCysAlAGluIlePheProThrArgValAlaArgIlyLeuCysIleAlaI	727
Db		84681	TCCCCAACATTGTGTGCGGAATCTTCCCACATAGGCGCCGCGGATCTGCATGGCA	84622
Oy		727	IeCysAlaPheThrPheThrIleGlyAspIleIleValIThryrSerLeuProValMetL	747
Db		84621	TCTGGCGCCCTGACATTTCTGATTTGGTATATCATAGTACACTTACAGCTCCCTGTATGC	84562
Oy		747	euaenAlaIleGlyLeuAlaGlyValPheSerIleTyralaValCysLeuIleSerp	767
Db		84561	TGAATGGCATGGGCTTGACAGGTCTCTTGGTANATACGAGTGTGTCGATTTGCCT	84502
Oy		767	heValPheValPheLeuLyrsValProGluThyrIysGlyMetProLeuGluValIleThrG	787
Db		84501	TCTGGTTCTGCTTCTCTCAAGGTCCCCGACGAGCAAGGCGATGCGCTGAAATCATCACCG	84442
Oy		787	IuPhePheAlaValAlGlyAlaIlysglnAlaIlaIalaIa	798
Db		84441	AGTTCTTGCTGTGTGTGTCANAAGCAATAATGCAGCT	84407
RESULT 4				
LOCUS	HVU534445			
DEFINITION	Hordeum vulgare mRNA for hexose transporter (stpl gene).			
VERSION	AJ534445.1 GI:26986185			
KEYWORDS	hexose transporter; stpl gene.			
SOURCE	Hordeum vulgare subsp. vulgare			
ORGANISM	Hordeum vulgare subsp. vulgare			
REFERENCE	Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
AUTHORS	Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;			
	Poideae; Triticeae; Hordeum.			
	1			
	Webesch,W., Pantlitz,R., Gubatz,S., Wang,Q., Radchuk,R., Weber,H. and			
	Wobus,U.			
TITLE	The role of invertases and hexose transporters in controlling sugar			
	ratios in maternal and filial tissues of barley caryopses during			
	early development			
JOURNAL	Plant J. 33 (2), 395-411 (2003)			
MEDLINE	22424051			
PUBMED	12535352			
REFERENCES	2 (bases 1 to 2614)			
AUTHORS	Radchuk,R.			
TITLE	Direct Submision			
JOURNAL	Submitted (09-DEC-2002) Radchuk R., Molecular Genetics - Gene			
	Expression, Plant Genetics and Crop Plant Research, Corrensstr. 3,			
	Gatersleben, D-06466, GERMANY			
FEATURES	Location/Qualifiers			
	1..2614			
	/organism="Hordeum vulgare subsp. vulgare"			
	/mol_type="mRNA"			
	/culivar="barke"			
	/db_xref="taxon:112509"			
	/tissue_type="developing caryopsis"			
	/clone_lib="lambda DAPF2"			

```

gene
CDS
    /dev stage="0-15 days after flowering"
    1..2614
    /gene="stpl"
    83..2314
    /gene="stpl"
    /codon_start=1
    /product="hexose transporter"
    /protein_id="CAD58958.1"
    /translation="MSSGAVLVVAISIGNLLQGM DNATLAGAVLYIKKERSLETOPIL
    EGIIVMSLIGATVITFSGVADVADEGRPLIASVYVSLMELMANVYLLMR
    ARLIDGGIAGLAVTVLPIVLTSETAPDRIIGRLNLTLPQSSGGMEISCMVFTSLMR
    QPDRIMLGVLSIPSLMTPALTYVLTPESRPLVSKRMBAKRVLORLKREVSSE
    MALLVEGLGKVDHFEFVEYIIGPDELADLDGALDQELKLYGEBEGVMIAPEVRGK
    GOSMISGALGMDHSHGEMVSGQKGLVPLVLTFSVSEHKPEVSNRSTLFPNFGSK
    FSVAGQOQAKADMDAESRDDEVDASDHQADDIEDNINSPLISQATSVBEKEIAAPGS
    GSIVGVESSSMOGGDAVSSGIGGQGMKMKWTERRGADQKGEGRVLYLHEGVS
    GDRKSLTSMPGDIPPGCEYIQAAALVSGPALYSKRLIQOLAGPMVHSEAVARG
    TKMAELFEPGVKALFVIGIQILIQDPAINGVLYTTPQIDBOAGVILISNLSGSG
    SASLISALTTLLMLPSIIGIAMRLMDSGRFFLLSTPVLVALAILVIVNLVDGSI
    MTHALMNSTIATVIECFVWGFPIIPILCAETPESTRGICLICALTFWIGIIVATY
    YTLPELMNAIGLAVGFIVAVVCMIAEFVYMKVPEPKMPLFVITFEFSSVGAQKGE
    ATD"
BASE COUNT      563 a      586 c      755 g      708 t      2 others
ORIGIN
Alignment Scores:
Pred. No.:      4,24e-178      Length:      2614
Score:          2808.00      Matches:      543
Percent Similarity: 84.39%      Conservative: 84
Best Local Similarity: 73.08%      Mismatches: 104
Query Match:    68.91%      Indels:      12
DB:             8      Gaps:        6

US-10-051-909-32 (1-800) x HVU534445 (1-2614)
QY      57  SerGlyAlaValLeuValAlaTlleValAlaSerIleGlyAsnLeuLeuGlnGlyTTPasp  76
Db      89  TCGGGCGCGCGTGGCTGCGCATAGCGGCGCTCCATCGGAACCTGCGACAGGGGTGGGAT  148
QY      77  AsnAlaThrIleAlaAlaAlaValLeuTyrIleLysIysGluPheGlnLeuGlnAsnGlu  96
Db      149 AATCGACCATTTGACAGGTGCGCTGCTGACATTAAGAAAGAGTTCCAGCTCGGAACCCAG  208
QY      97  ProThrValGlnGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrThr  116
Db      209  CCTTTGATCGAGGCGCTCATGTCGTCATGCGTCCATCGAGGACGAGGTATACACAGC  268
QY      117 PheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSerIle  136
Db      269  TTCTCCGGGGCGGAGGTGACGCTGTTGGTGAAGGCGGCCCTGTGATGCGCTCGTGTGC  328
QY      137 LeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeuLeu  156
Db      329  CTCAACTTTCACATGCGCTGATGATGCTCGGGCGCCCAACGCTCATGTTGCTCTTG  388
QY      157 AlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIle  176
Db      389  GCCAGGCTCATCGACGGGTTGGATACGATTGGGACGACCTTGTCTCTCTGTACATT  448
QY      177 SerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGly  196
Db      449  TCGGAGACCGCCCGACATGACATTAGAGGGCTGCTGAACACAGCTGCCGCAATTAGTGG  508
QY      197 SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSerPro  216
Db      509  TCAGAGAGGAGTTCCTTTCTTACATGATGATGTTTACCATGCTCCCTCATGCCACAGCT  568
QY      217 AspTPArgIleMetLeuGlyValLeuAlaIleProSerLeuAsnPhePheGlyLeuThr  236
Db      569  GACGTGAGAACATCAGCTTGGGGGTTTTTGTGCATCCCGTCGCTTAATGATTTTGCATTGACT  628
QY      237 IlePheTyrLeuProGlnIleProArgTyrIleValSerIleLysGlyArgMetAlaGluAla  256

```

```

Db 629 GCTTTTATTTCCCGAGTCGCCAGATGGCTTTGAGCAAGGAAAGATGGCTGAGGCC 688
Qy LysIlyValLeuGlnIlyLeuArgIlyAspAspValSerGlyGluLeuSerLeu 276
Db 689 AAGCAGATTTCCAAAGACTCGCGGGAGGAAAGATGCTCAAGAGAGATGGCCCTTCTT 748
Qy LeuGlnIlyLeuGlnValGlyIlyAspThrSerIleGlnIlyTrpIleIleGlyProAla 296
Db 749 GTTGAAGAGATGGGTGTGGAAAGACACATTTTGGAGAAATACATAATTGGACCTGAT 808
Qy ThrIleAlaIleAspAspLeuValThrAspGlyAspIlyGlnIlyGlnIleThrIlyGly 316
Db 809 GATGAGCTGGCTGATGACGCGGCTGCTCA---GATCAAGAGAAAGTTGAACTTATGGA 865
Qy ProGlnIlyGlnIleSerTrpIleAlaArgProSerIly-----GlyProIleMetLeu 334
Db 866 GCTGAAGAAAGGGATATTCTTGATCGCCGCTCTGTAGSGGTGGCGCCAAAGTGCACTT 925
Qy GlySerValLeuSerLeuAlaSerArgIlyGlySerMetValAsnGlnSerValProLeu 354
Db 926 GGAACCGCTTTGGGCTCATGTCTGTCATGGAGATGTTAGTTCAGGGTAAATCTCTT 985
Qy MetAspProIleValThrIleuPheGlySerValIleAsnMetProGlnIleGlyIly 374
Db 986 GTTGAACCCCTTGCTCATCTCTTGGAGATGTCATGAGAAAGATGCTTAGGTAATGGGG 1045
Qy SerMetArgSerThrIleuPheProAsnPheGlySerMetPheSerValThrAspGlnIle 394
Db 1046 AGCATGCGGAGACCTGTTTCTTAATTGGCAGATGTTTAGTGGCTGTAACAGCAG 1105
Qy AlaIlyAsnGlnIlyTrpAspGlnIlyuLeuIleAsnIlyAspAspGlnIlyTrpAlaSer 414
Db 1106 CAGGTTAAAGCTGACTGGATGCTGAGAGT---CATAGGATGATGAAATATATCATCTG 1162
Qy AspGlyAlaGlyIlyAspTrpGlnAspAsnLeuIleSerProLeuSerArgGlnIle 434
Db 1163 GATCATGGCTGATGACATTTAGAGATACCTCAATAGCCCTTATTTCTGCTCAAGC 1222
Qy ThrGlyAlaGlnIlyLysAspIleVal-----HisIleGlyHisArgIlySerAlaLeu 452
Db 1223 ACAACGCGAGAGGATGAGATGCTGCACTCAAGAAAGATAGGTGGTGGGA 1282
Qy SerMetArgArgGlnThrIleuLeuGlyIlyGlyIlyAspGlyValSerSerThrAspIle 472
Db 1283 AGTACTAGCATGAG-----GGAGGGAGCGCATGACGACGATGGGCAAT 1327
Qy GlyIlyGlyTrpGlnLeuAlaTrpIlySerTrpSerGlnIlyGlnIlyGlnAsnGlyArgIly 492
Db 1328 GGTGGGGGGGTGAGATTAGCTTGAAGTGAAGAGAGAGGTGCGAGATGGGCAAAA 1387
Qy GlnIlyGlyPheLysArgValIlyTrpIleuHisGlnIlyGlyValProGlySerArgArgIly 512
Db 1388 GAAGGTGGCTTCCAGCATATTACTTGCATGAGAGGGGTGTCAGGTATCGGAGAGCC 1447
Qy SerIleValSerLeuProGlyIlyGlyIlyAspValPheGlnIlySerGlnPheValHisAla 532
Db 1448 TCTATATTGTCTATATGCA---GGAGGTGATATTCTCCGCGGTGAGATATCCAGGCA 1504
Qy AlaAlaIleuValSerGlnSerAlaLeuPheSerIlyGlyIlyGlnIlyGlnProArgMetSer 552
Db 1505 GCGGCTCATGATGACCACTGCTCTTATCTGAAGAGACATATATAGCAGCAGCTGCT 1564
Qy AspAlaAlaMetValHisProSerGlnValAlaAlaIlyGlySerArgTrpIlyAspLeu 572
Db 1565 GGTCAAGCCATGATCATCATCCGAGCGAGTGGCAAGGGATCAAAAGTGGGCAAGACTA 1624
Qy PheGlnProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnIly 592
Db 1625 TTTGAACCTGAGATGAGATGACACTGTTGTTGGATGATGATTAAGATCTCGCAGCAG 1684
Qy PheAlaGlyIleAsnGlyValLeuValTrpTrpProGlnIleLeuGlnIleAlaGlyVal 612

```

```

Db 1685 TTTGGGGTATCAACGAGTCTCTGATATTACAGCCTCAGATCTTGAAGACAGTGT 1744
Qy AlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSer 632
Db 1745 GGGATTTCTTATATCAAAATTTGAGACTTAAGCTTCTCTGATCATATCTTATTAAGTCC 1804
Qy LeuThrTrpIleuLeuMetLeuProCysIleIlePheAlaMetLeuLeuMetAspLeuSer 652
Db 1805 TTGACACCTTGCTGATGCTTCCAGCATTTGGCATGCGCATGAGACTCATATGATGTCA 1864
Qy GlyArgArgPheLeuLeuLeuGlyIlyThrIleProIleLeuIleAlaSerLeuValIleLeu 672
Db 1865 GGAAGAGATTTCTTCTTCTCTCAACATCCCTGCTTGATAGTGCATGATGATCTTGG 1924
Qy ValValSerAsnLeuIleAspLeuGlyIlyThrLeuAlaHisAlaLeuLeuSerThrValSer 692
Db 1925 GTTTGGTGAATGTTCTGATGATGAGAAACATGTGGACGCGGCTCTCTCAACGATCAGT 1984
Qy ValIleValIlyTrpPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCys 712
Db 1985 GTCATAGTCTACTCTGCTTCTGTCATGGGTTTGGACCTATCCGAAATATCTCTGC 2044
Qy AlaGlnIlePheProThrArgValArgIlyLeuCysIleAlaIleCysAlaPheThrPhe 732
Db 2045 GCGAGATTTTCCCACTCCGTCGTCGATGATGATGATGATGATGATGATGATGATGATG 2104
Qy TrpIleGlyAspIleIleValIleThrIlySerLeuProValMetLeuAsnAlaIleGlyLeu 752
Db 2105 TGATGTTGATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2164
Qy AlaGlyValPheSerIleIleValIleValIleCysLeuIleSerPheValPheValPheLeu 772
Db 2165 GCTGAGTTTCTGGAATATATGCTGCTTGTGATGATGATGATGATGATGATGATGATGATG 2224
Qy LysValProGlnTrpIlyGlyIlyMetProLeuGlnValIleThrGlnPhePheAlaValGly 792
Db 2225 AAGGCTCTGAGACAAAGGAGATGCTGAGAGTCAACGAGATCTTCTCGGTGGG 2284
Qy 793 AlaIlyGln 795
Db 2285 GCGAAGCAG 2293
RESUT 5
AR208568 LOCUS 2601 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 7 from patent US 6383776.
ACCESSION AR208568
VERSION AR208568.1 GI:21509755
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2601)
AUTHORS Allen, S.M., Hitz, W.D., Kinney, A.J. and Tingey, S.V.
TITLE Plant sugar transport proteins
JOURNAL Patent: US 6383776-A 7 07-MAY-2002;
FEATURES
source location/Qualifiers
1..2601
/organism="unknown"
BASE COUNT 648 a 543 c 644 g 766 t
ORIGIN
Alignment Scores:
Pred. No.: 3,74e-169 Length: 2601
Score: 2674.00 Matches: 522
Percent Similarity: 82.20% Conservative: 92
Best Local Similarity: 69.88% Mismatches: 115
Query Match: 65.62% Indels: 18
DB: Gaps: 8
US-10-051-909-32 (1-800) x AR208568 (1-2601)
Qy 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnIlyTrp 75

```

Db 175 ATGAAAGGTGCGCTTGTGCTATTCGCGCTTCATTCGTAATTCCTCAAGAGATG 234
 Qy 76 AspaenalaThrIleAlaIleValLeuTyrIleuIleuPheGlnLeuIleuIleu 95
 Db 235 GATTAATGCTACCATGCGCGGCTTAAGTATTAAGAAAGACCTTCTTGGGAAACA 294
 Qy 96 GluProThrValIleuIleValSerMetSerLeuIleValAlaThrIleValThr 115
 Db 295 -----ACTATGGAAAGGCTTGTGCTGATCTCCCTGATTTGGAGCAACGGTAATCACC 348
 Qy 116 ThrPheSerGlyProLeuSerIleGlyIleValArgProMetLeuIleuSerSer 135
 Db 349 ACATGCTGCTGCTATAGCGATTTGCTCGGTCGCGACCATATGATATCTCATCT 408
 Qy 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrPheProAsnValTyrValLeu 155
 Db 409 GTGCTCTATTTCTTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 468
 Qy 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
 Db 469 TTGGGAGGCTATCTGATGATGATTTGGGATTTGGGCTTGGCTGATCTTGTCCGCTCAT 528
 Qy 176 IleSerGlyIleAlaProSerGlyIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195
 Db 529 ATATCTGAAGCGCGCGCTGAATTAAGGGGCTGGTGAATACGCTTCTCATGTTCACT 588
 Qy 196 GlySerGlyIleMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
 Db 589 GGCTGTGAGGAATGTTTGTGCTATGATGATGATGATGATGATGATGATGATGATG 648
 Qy 216 ProAspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
 Db 649 CCTACTGAGGCTATGCTTGGGCTTGTGCTATTTCTTCTTCTTGTATTTGCTATG 708
 Qy 236 ThrIlePheTyrLeuProGlyIleProArgTyrLeuValSerIleGlyIleArgMetAlaGly 255
 Db 709 ACCATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 768
 Qy 256 AlaIleValValLeuGlyIleValLeuArgGlyIleValAspAspValSerGlyIleLeuSerLeu 275
 Db 769 GCTAAGAGAGTCTCCAAAGATTTGGCGGAGAGAGATGTCAGGCGAGATGTCATG 828
 Qy 276 LeuLeuGlyIleValLeuValIleGlyIleAspThrSerIleGlyIleTyrIleIleGlyPro 295
 Db 829 CTGGTTGAAGTCTCGGATTTGGGCTGATACATCTATCAAGAGATCAATTTGACCT 888
 Qy 296 AlaThrGlyAlaIleAspAspLeuValThrAspGlyAspIleGlyIleThrIleTyr 315
 Db 889 GCTGACATGTCGCTGATGCTGATGACATGACAGACAGAGAAATGATTAAT 948
 Qy 316 GlyProGlyIleGlyIleSerTyrIleAlaArgProSerGlyIleProIleMetLeuGly 335
 Db 949 GGATCCCAAGCAGGCTTCTTGTGATCAAAACCTGTCACAGAGTTCTATTTGGC 1008
 Qy 336 SerValIleSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProMet 355
 Db 1009 -----CTTGGCTCAACCATGAGACATCATCAACCAAGCATGCCCCCTCATG 1056
 Qy 356 AspProIleValThrLeuPheGlySerValHisGlyIleMetProGlnIleGly----- 373
 Db 1057 GATCTCTGCTGACATGCTTTGGTAGCATTCATGAGAACTCCCGACAGACAGACAGA 1116
 Qy 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
 Db 1117 GGAAGCATCGAAGCATCTGTTTCCAAATTTTGGAAAGCATGTTGACGACCTGTCGCGCG 1176
 Qy 394 HisAlaValAsnGlyIleThrAspGlyIleValAsnLeuHisArgAspAspValGlyIleVal 413
 Db 1177 CATGCTAAATTTGAACATGAGATGAGAAAGCTTCAAAAGGAGAGTGAAGATCATG 1236
 Qy 414 SerAspGlyAlaGlyIleAspTyrGlyIleAspAsnLeuHisSerProLeuLeuSerArgGln 433

Db 1237 TCAGATGCACCCGTTGGGACCTCCGATGATTAATTGCACAGTCTTAACTCAGCCAA 1296
 Qy 434 AlaThrGlyAlaGlyIleValAspIleValHisGlyIleHisArgGlySerAlaLeu--- 452
 Db 1297 ACAACAGCCTTGA---AAGACTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1353
 Qy 453 SerMetArgArgGlnThrLeuLeuGlyIleGlyIle---GlyAspGlyValSerSerThrAsp 471
 Db 1354 AGCATGAGGCTCAACAGTATCTCATGCAAGGTCAGGTGAGCAAGGTGATGATCAGCT 1413
 Qy 472 IleGlyIleGlyIleTyrPheIleuAlaTyrPheTyrPheSerGlyIleGlyIleValArg 491
 Db 1414 ATTTGCTGTGCTGCTGCAACTGCGATGAGAAATGACTGTAA---GGTGAAGATGAA 1470
 Qy 492 LysGlyIleGlyIlePheValArgValTyrLeuHisGlnGlyIleValProGlySerArgArg 511
 Db 1471 CAACAGAGAGGCTTAAAGATTTATTAATAGAGAGGAGATTTCTGCATCTCGTCTG 1530
 Qy 512 GlySerIleValSerLeuProGlyIleGlyIleValAspValPheGlyIleSerGlyIleValHis 531
 Db 1531 GGATCCATTTGATGATTTCCCGTGAAGC-----GAATTTGTCCAG 1572
 Qy 532 AlaAlaIleValSerGlnSerAlaLeuPheSerIleGlyLeuAlaGlyProArgMet 551
 Db 1573 GCTGCTGCTGCTGTAAGCCACCGCTCTTACTCAAGAGCTTATGATGACACCCCA 1632
 Qy 552 SerAspAlaIleMetValHisProSerGlyIleValAlaIleValSerArgTyrIleAsp 571
 Db 1633 GTTGGGCTGCAATGTTGCTACCCATCTGAGACACCTTCAAGGGGCGCAAGTTGGAACCT 1692
 Qy 572 LeuPheGlyProGlyIleValArgAlaLeuLeuValGlyIleGlyIleGlnIleuGln 591
 Db 1693 CTTCTGACACAGGCTTACACATGATGCTGTTGAGTGTGAGATGCAATCAATCTTACG 1752
 Qy 592 GlnPheAlaGlyIleAsnGlyIleValLeuTyrTyrThrProGlnIleLeuGlyIleAlaGly 611
 Db 1753 CAGTTTCAGGATTAAGAGGCTTCTATTAACACACTCAAACTTGAAGAGCGCGCT 1812
 Qy 612 ValAlaValIleLeuSerIlePheGlyIleuSerSerAlaSerAlaSerIleLeuIleSer 631
 Db 1813 GTTGAAGTCTTCTTCAATTAAGCAATGCTGCTGAGAGTGCATCTTCTATATCGT 1872
 Qy 632 SerLeuThrThrIleLeuMetIleuProCysIleGlyPheAlaMetLeuMetAspLeu 651
 Db 1873 GCTTTCACACACTTCTTATGCTTCTCTGATTAAGCGTACAGCAATGATGATGAT 1932
 Qy 652 SerGlyArgArgPheLeuLeuGlyIleThrIleProIleLeuIleAlaSerLeuValIle 671
 Db 1933 TCAGGCAAGAGCAGTCTCTACTTACTTACATCCCGGTGATGTTGCTCATTTAT 1992
 Qy 672 LeuValIleSerAsnLeuIleAspLeuGlyIleThrIleAlaHisAlaLeuSerThrVal 691
 Db 1993 TTGGCATTTGAAGCTGTGTAATTTTGGCAATGTCGCCATGCAAGATCTCAACAGTA 2052
 Qy 692 SerValIleValTyrPheCysCysPheValIleMetGlyPheGlyProIleProAsnIleLeu 711
 Db 2053 TGCCTTGTGATTTATTTCTGCTGCTTGTGATGAGGTATGAGCAATTCCAACATCTCT 2112
 Qy 712 CysAlaGlyIlePheProThrArgValArgGlyIleCysAlaIleAlaIleCysAlaPheThr 731
 Db 2113 TGCTCAGAGATTTTCCACATAGGCTGCTGCGCTCATATGCTGATGATGATG 2172
 Qy 732 PheTyrIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGly 751
 Db 2173 TTCTGATTTGAGACATCATCATCAATCTGCTGCTGATGATGCTGCTGCTTATGAG 2232
 Qy 752 LeuAlaGlyValPheSerIleTyrAlaIleValCysLeuIleSerPheValPheValPhe 771
 Db 2233 CTGCTGCTGATTTGCTCATTTACGCACTTGTGTTCTTCTCATCTGCGATTTGTGTT 2292
 Qy 772 LeuIleValProGlyIleThrIleGlyIleMetProLeuGlyIleIleThrGlnPhePheAlaVal 791
 Db 2293 TTGAAGTTCAGAAACAAAGGAGTGCCTTGAAGTCACTCTGATTTCTTCTCTGTT 2352

QY	792	G V A L A T G G C A G C G C T G C T T C T	798
Dd	2353	GGAGCAACGACGCCTGCTTCCT	2373
RESULT 6			
LOCUS	AY094465		
DEFINITION	Arabidopsis thaliana ATg35300/F23E12_140 mRNA, complete cde.		
ACCESSION	AY094465.1	GI:20453188	
VERSION			
KEYWORDS	FLI CDNA.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids; euroside II; Brassicales; Brassicaceae; Arabidopsi		
REFERENCE	1 (bases 1 to 2570)		
AUTHORS	Shimn, P., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Banh,J., Bower,L., Carinci,P., Chang,B., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.		
TITLE	Arabidopsis cDNA clones		
JOURNAL	Unpublished		
REFERENCES	2. (bases 1 to 2570) Shimn, P., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Banh,J., Bower,L., Carinci,P., Chang,B., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R. Direct Submission Submitted (08-APR-2002) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pine Road, La Jolla, CA 92037, USA RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carinci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.		
COMMENT			
TITLE	The salk, stanford, pgsc (ssp) consortium members carried out the sequencing and annotation of the rafL cDNAs: Shimn,P., Chen,H., Cheuk,R., Kim,C.J., Meyer,M.C., Banh,J., Bower,L., Chang,B., Dale,J.M., Goldsmith,A.D., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Davis,R.W., Theologis,A., and Ecker,J.R.		
FEATURES			
SOURCE	Shimn,P. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs. Location/Qualifiers 1..2570 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /chromosome="4" /clone="RAFLO9-48-J19(R21968)" /note="ecotype: Columbia" 209..2428 /note="putative sugar transporter protein" /codon_start=1 /product="ATG435300/F23E12_140" /protein_id="AAIM19835.1"		
CDS			

		/db xref="GI:20453189"			
		/translation="MSGAVIVA1AANVCNLLQGDMDNTTGAVALYIKENPLKSNPS			
		EGIVAMSLIGLTLITTSQGVADMLGRPMILISTILYFGSLWMLSNVYVLL			
		RLDVGGLVTVTLPIYISETAPEELRGLNLTLPQFTSGGMFLSTCYMFGMSLMP			
		PSRMILGVLPIPSLVFFPLVYFIPESPRMLVSKGMLAKVLOLRORREAVSG			
		ALLVEGLIGSTTIEEYIIGPADVTDHDIADVKKQIKLYGAEGLSVAVPVKG			
		STMSVLSRHSSTMSRQGLIDPLVTLTGSHKEMPTQMSRSLPFPQSMVSYS			
		ORHREMDENLVGSGEDYPSDHPDSDDLASLISQRTMSMCPHRAHDTSLP			
		RHSQYQAGQEGEASMSKIGGVOYAKMTTEREDSSQKSGFRKITYLDEGPGSR			
		GSIVLPFGDGTGEADPVQASALVSQPLVSKDILKHTTIGPAVWVSETTKS			
		LHDPEKRALVVGVLQILQOFSGINLVTYTPDILQDAQGLISNMGISSSASIS			
		IALTFPMLPAIAVAMRLMLSGRRTLTLPILASLIVLIVISLVHMSIVHA			
		LSVSLVLYFCFPMVGFAPANILCSLIFPRVAGICIALCALFWICDIIVYSLP			
		LKSQIGAGVFCFMAIVVCCISMWPIKVPETKMPLEVITPEFSVQAQAEAKNIE			
BASE COUNT	629 a	500 c	646 g	795 t	
ORIGIN					
Alignment Scores:					
Align. No.:	5, 02e-168	Length:	2570		
Score:	2657.00	Matches:	530		
Percent Similarity:	79.87%	Conservative:	97		
Best Local Similarity:	67.52%	Mismatches:	133		
Query Match:	65.20%	Indels:	25		
DB:	8	Gaps:	10		
US-10-051-909-32 (1-800) x AY094465 (1-2570)					
Qy	29	LeuAlaLeuProGlyProLeuProProAlaSerCysSerSerGinglu-----	44		
Db	91	CTTCTCTGCGCT---CNAFTTCAGAGGATTTGTTTCACTACCTACGTCGTTGATC	147		
Qy	45	-----ProValIhSerAspApIleLeuGluAlaPlyS-----	55		
Db	148	AGTACCTGTTTTGCGAATTCTCTTCAATTTCTTGATTAAGATAGACGGAAT	207		
Qy	56	-MetSerGlyAlaValleuValAlaIleValAlaSerIleGlyValMetLeuGingluYTr	75		
Db	208	TATGAGTGAGCGTGGCTTTGTCATTGCTGCTGCTGTGGCACTTGTTACAAGATG	267		
Qy	75	PAEPaenAlaThrIleAlaAlaValaValLeuYrIleValSylGluPheGluLeuGlnAs	95		
Db	268	GGATACGCACTATTGCGAGAGCGTGTGTGTACATAAAGAGTTTATTGGAGAG	327		
Qy	95	ngluProthrValGluGlyLeuIleValaSerMetSerLeuIleGlyAlaThrIleValTh	115		
Db	328	TAACTCATGAGGGAAGGCTTAATTGTGGCAATGTCATTAATTTGCTACTCGATTAC	387		
Qy	115	rThrPheSerGlyProLeuSerAspSerIleGlyAlaArgProMetLeuIleLeuSerSe	135		
Db	388	AACATGCTCGAGGGGAGTGAATGAGCTGTGGCCGTGCCATGCTAATATTGTCTC	447		
Qy	135	rIleLeuYrPhePheSerGlyLeuIleMetLeuTrpSerProAlaValTyValleuLe	155		
Db	448	AATTCTCTACTTTGTTGGTTCCTAGTAATGCTAATGCTCCGAATGTTATGTGTTGCT	507		
Qy	155	uLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTy	175		
Db	508	CTTAGAAGAGTTGTTAATGATGATTTGGGGTTGGTCTTGCGTCAACTTGTCTTATTA	567		
Qy	175	rIleSerGluIleAlaProSerGluIleArgGlyLeuMetAsnThrLeuProGlnPheSe	195		
Db	568	TAAATCTGAGCTGACCACTGGAATTAAGGAGCTGTGAATVACGCTACCGCAGTTTAC	627		
Qy	195	rgIleSerGlyGlyMetPheLeuSerTyCysMetValPheGlyMetSerLeuSerProSe	215		
Db	628	TGGCTCTGAGGGAATGTTCTTATCTTACTGTATAGTGTTCGGAATGTCGTTAGTCATC	687		
Qy	215	rProAspTrpArgIleMetLeuGlyValleuAlaIleProSerLeuPhePhePheGlyLe	235		
Db	688	ACCTACTGGAATGATGAGCTTGGTGTCTTTCATCCCTTCCTGTCTTTTCTTCTCT	747		
Qy	235	uThrIlePheTyrLeuProGluSerProArgTrpLeuValSerIleGlyArgMetAlaGl	255		

Db 748 CACGGCTCTTCTTCCGCCAGTCCCAAGTGGCTCGTAGCAAGTGCATGCTTGA 807
 Qy 255 uAlaIyValIleuGlnIleuArgIlyAspAspValIleuSerGlyIleuSerIe 275
 Db 808 AACCAAGCGGATTCCTTCAAGAGCTGCTGCGCAAGATGTCGCGAGATGGCTT 867
 Qy 275 uLeuLeuGlnIyLeuGlnValGlyIyAspThrSerIleGlnIyIleGlyPr 295
 Db 868 GTTGGTGAAGGCTTCGAAATTGAGAGTGAACAACATAGAGATATATATATGCTTC 927
 Qy 295 oAlaThrGlnValaAlaAspAspLeuValIThrAspGlyAspIyGlnIleThrLeuTy 315
 Db 928 CCCGATGAGTACTGATGATCATGATATAGCTGTGATTAAGATCAATTAAGTTATA 987
 Qy 315 rGlyProGlnIyGlnIyGlnSerTrpIleAlaArgProSerIyGlyProIleMetLeuG 335
 Db 988 TGGTCAGAGAGAGGCTGAGTGGGTCTGAGGCCAGTCAAGCAAGG-----GG 1035
 Qy 335 ySerValIleuSerIleuAlaSerArgIleGlySer---MetValAsnGlnSerValProIe 354
 Db 1036 AAGCATAGATAGATGTTTGTCTCGCATGAGATACAAATGACAGAGGCAAGCTCAT 1095
 Qy 354 uMetAspProIleValIThrLeuPheGlySerValIleGlnAsnMetProGlnAlaGly 374
 Db 1096 GATTGATCCTCTGTCACACTGTTTGGAGCGCTTCAGAGAAAGATCCGAGACACT---CG 1152
 Qy 374 ySerMetArgSerTrpIleuPheProAsnIleGlySerMetPheSerValIThrAspGlnI 394
 Db 1153 AAGCATAGAGAGTCCCTTGTCTCCACATTTTGGAGATGTTTCAGTGTGAGGGAATCA 1212
 Qy 394 sAlaIyAsnGlnIyGlnIyAspGlnIyAsnLeuIleAspAspAspGlnIyIleAla 414
 Db 1213 ACCAAGACATGAGATGGAGATGAGAAAGATCTTGTGAGAAAGGAGAGATATTCATC 1272
 Qy 414 rAspGlnAlaGlyIyAspIyGlnIyAspAsnLeuIleSerProLeuLeuSerArgGlnAl 434
 Db 1272 CGAAC---CATGGAGATGATTCGAGATGATCTTCATCTCCGTTGATCTCAGCTCAAC 1329
 Qy 434 aThrGlnAlaGlnIyGlnIyAspIleValIleIleGlnIleGlnIyIleAlaLeuSerMe 454
 Db 1330 GACAGAGCATGAG---AAAGACATGCTCACAACCTGCTCATGAACTTTTCACTTCAG 1386
 Qy 454 rArgArgGlnTrpIleuLeuGlnIyGlnIyAspGlnIyAspSerThrAspIleGlyI 474
 Db 1387 AATATGAGATCAAGTCAAGGAGCTCAAGGAGGAGGAGGAGGAGATGATGAGGATTCGAG 1446
 Qy 474 yGlyTrpGlnLeuAlaTrpIyIleTrpSerGlnIyGlnIyGlnIyAsnGlnIyArgIyGln 494
 Db 1447 TGGATGGCAAGTGGCATGGAATGAGCGAAAGAGATGATGATCGGACAGAAAGAGAG 1506
 Qy 494 yGlyPheIyValGlnValIyIleuIleGlnIyGlnIyValProGlnIySerArgIySerI 514
 Db 1507 TGGGTTTAAACGATATCTTCATCAGAAAGGTTTCCAGAGATTCGAGCTGGCTCAT 1566
 Qy 514 eValSerIleuProGlnIyGlnIyAspValPheGlnIyGlnIyGlnIyIleAlaAla 534
 Db 1567 TGTTCATTCCTCGTGGTGGATGAGAAACCGGTGAG---GAGATTTTGTCAAGCGCTCC 1623
 Qy 534 aLeuValSerGlnSerIleuPheSerIyGlnIyLeuAlaGlnProArgMetSerAspAl 554
 Db 1624 TTTGGTTAGCAACGAGCTTTATTCAAAGACCTTCCAAGAACATACATACATTCGTC 1683
 Qy 554 aAlaMetValIlePheProSerGlnValAlaAlaIyGlnIySerArgIyTrpIyAspLeuPheG 574
 Db 1684 TGTCTAGTACATCCATCCGAA---ACAACTAAAGGTCATATTTGCAATGATCTTCATCA 1740
 Qy 574 uProGlnIyValArgArgAlaLeuLeuValGlnIyGlnIleGlnIleuGlnIlePheAl 594
 Db 1741 TCTGTAGTCAACGCGCTTACTGATGAGATTCGATTCATATCTTCAGCAGTTCTC 1800
 Qy 594 aGlyIleAsnGlnIyValIleuTyTrpThrProGlnIleuGlnIyGlnIyValAlaIyAla 614
 Db 1801 AAGCATCAACGAGGTTCTTACTACACACGCAATCTTTGAGCAGGCGGCTGCGGAT 1860

Qy 614 lIleuSerIyPheGlyIleuSerSerAlaSerIleuIleuIleuSerSerIeTh 634
 Db 1861 CCTACTATCGAACATAGGAGATTAAGTTCTTCTCCAGCATCTTACTTATTAAGTGCATTAC 1920
 Qy 634 rThrLeuMetIleuProCysIleGlyPheAlaMetLeuMetAspLeuSerGlyIyAr 654
 Db 1921 AACCTTGTGATGTTACTCGAAATAGCTGTTCATATGAGGCTCATGATGATCTTTCGTG 1980
 Qy 654 gArgPheLeuLeuGlnIyThrIleProIleuIleuAlaSerIleuValIleuValIy 674
 Db 1981 AAGGACCTTGTCTTCACACAGATACCAATCTGATACATCTTATGATGTTTAT 2040
 Qy 674 lSerAsnIleuLeuGlnIyThrLeuAlaIleAlaLeuSerThrValIleValI 694
 Db 2041 CTCAATCTTGTTCACATGAACAGATTTGTCAGCGGCTTATATCAACGTAAGCGTTGT 2100
 Qy 694 eValTyPheCysCysPheValMetGlyPheGlyProIleProAsnIleuLeuCyAlaG 714
 Db 2101 GCTTACTTGTCTTCTTGTGATGAGTGGTTCGTCTGCTCCAAACATCTCTGTTCAGA 2160
 Qy 714 uIlePheProThrArgValArgIyLeuCysIleAlaIleCysAlaPheThrPheTrpI 734
 Db 2161 GATTTTTCACATCCAGTCCCGGAAATCGATCGCATCTGCGCACTCACCTTCTGAT 2220
 Qy 734 eGlyAspIleIleValIThrTySerIleuProValMetLeuAsnAlaIleGlyLeuAla 754
 Db 2221 CTGTACATTAATCGTCACTTACAGTCTCCGCTGCTCAAAATCATTTGATGATGCTGG 2280
 Qy 754 yValPheSerIleTyAlaValAlaValCysLeuIleSerPheValPheValPheLeuIyVa 774
 Db 2281 TGTGTTGGATGATGATCGCAATCTGATGATTCATTCAGGAGCTTGTGTTCAATTAAGT 2340
 Qy 774 lProGlnIyIyGlnIyMetProLeuGlnIyValIleThrGlnPhePheAlaValGlyAla 794
 Db 2341 CCCGAAATCAAGGATGAGCATGAGTCAATGATCATCAGAGTCTTCTGTTGAGAGTAG 2400
 Qy 794 eGlnIleAlaAla 798
 Db 2401 ACAAGCTGAAGCT 2413

RESULT 7
 AX506620 2190 bp DNA linear PAT 27-SEP-2002
 LOCUS AX506620
 DEFINITION Sequence 1315 from Patent WO0216655.
 ACCESSION AX506620
 VERSION AX506620.1 GI:23387857
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1
 REFERENCE
 AUTHORS Harper J.F., Krebs J., Wang X. and Zhu T.
 TITLES Stress-regulated genes of plants, transgenic plants containing
 same, and methods of use
 JOURNAL Patent: WO 0216655-A 1315 28-FEB-2002;
 The Scripps Research Institute (US) ; Syngenta Participations AG
 (CH)
 FEATURES
 source location/Qualifiers
 1. 2190
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /db_xref="taxon:3702"
 BASE COUNT 529 a 439 c 572 g 650 t
 ORIGIN
 Alignment Scores: 2.99e-163 Length: 2190
 Pred. No.: 2584.00 Matches: 512
 Score: 2584.00 Conservative: 89
 Percent Similarity: 80.78%
 Best Local Similarity: 68.82% Mismatches: 123


```

OY 415 AaPgyAlaGlyGlyAspTyrGluAspAlaLeuHisSerProLeuLeuSerArgGlnAla 434
DB 1066 GAG---CATGAGAGATGATCTGAAGATGATCTTCATCTCCGTGATCTCAAGTCAAG 1122
OY 435 ThrGlyAlaGluGlyLeuAspIleValHisGlyHisArgGlySerAlaLeuSerMet 454
DB 1123 ACAACCATGAG---AAGACATGCTCTCACTGCTCAAGAACTCTTCTCACTTCACA 1179
OY 455 ArgArgGlnThrLeuLeuGlyGluGlyAspGlyValSerSerThrAspIleGlyGly 474
DB 1180 CATGAGAGCATGAGTGAAGAGAGCTCAAGAGAGAGAGCGGGTGTATGTGGGATTTGAGGT 1239
OY 475 GlyTyrGlnLeuAlaTyrPheTyrPheSerGlyGlyGluGlyAlaGlyValArgGlyGlyGly 494
DB 1240 GGATGGCAAGTGCATGAGAAATGAGACCGAAGAGAGAGATGAAATCGGAGACAGAAAGA--- 1296
OY 495 GlyPheLeuValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIle 514
DB 1297 -----GAGGTTTTCCAGAGATCTCGACGTGCTCAAT 1329
OY 515 ValSerLeuProGlyGlyGlyAspValPheGlyGlySerGlyPheValHisAlaAla 534
DB 1330 GTTCAATGCTGCTGGTGTGATGAGAACCGGTGAG---GCAGATTTTGTACAAACGCTGCT 1386
OY 535 LeuValSerGlnSerAlaLeuPheSerGlyGlyLeuAlaGluProArgMetSerAspAla 554
DB 1387 TTGGTTAGCCACCAAGCTCTTATCTCAAGACCTTCTCAAGAACATCATATGCTCT 1446
OY 555 AlaMetValHisProSerGlyValAlaAlaGlySerArgTyrPheAspLeuPheGlu 574
DB 1447 GCTATGATCATCATCCGAA---ACAACATAAGGGTCAATTTGGCATATCTTCATGAT 1503
OY 575 ProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnIlePheAla 594
DB 1504 CTTGAGATGACGCGATGATGATGAGAGAGTGTGACTTCAATCTTCAAGAGTTCCTCA 1563
OY 595 GlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaVal 614
DB 1564 GGCATCAACGAGGCTTCTTACTACACACCGCAATCTTGACACGGCGGTCTCGGAGATC 1623
OY 615 IleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThr 634
DB 1624 CTACATCATCAACATGCGGATGATGATCTTCTCAAGCATCTTCTTATTAAGTCAATGACA 1683
OY 635 ThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg 654
DB 1684 ACCTTGTGATGTACCTGCAATGATGCTGTGCAATGAGCTCATGATCTTCTGCTGCA 1743
OY 655 ArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValVal 674
DB 1744 AGGACCTGCTCTTCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1803
OY 675 SerAsnLeuIleAspLeuGlyThrIleAlaHisAlaLeuLeuSerThrValSerValIle 694
DB 1804 TCAATCTGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1863
OY 695 ValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuGlyValGlu 714
DB 1864 CTCCTCTGCTTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1923
OY 715 IlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIle 734
DB 1924 ATTTCCTCACTGAGTCCGCGAATCTGATGATGATGATGATGATGATGATGATGATGATGAT 1983
OY 735 GlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGly 754
DB 1984 TGTGACATATCGTCACTTACAGTCCCGGCTGCTCAATTCATTCATTCATTCATTCATTCAT 2043
OY 755 ValPheSerIleTyrAlaAlaValCysLeuIleSerPheValPheValPheLeuValVal 774
DB 2044 GTGTTTGAATGTAGCATATCGTATGTTGATTCATTCATTCATTCATTCATTCATTCATTCAT 2103
OY 775 ProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAlaLys 794

```

```

DB 2104 CCGGAACATAAGGATCCACTGAGATCATCAAGAGTCTTCTTCTGTTGAGACTAGA 2163
OY 795 GlnAlaAlaAla 798
DB 2164 CAAGCTGAAGCT 2175

RESULT 9
AC073166
LOCUS
DEFINITION
AC073166 142114 bp DNA linear PLN 28-AUG-2001
Oryza sativa chromosome 10 BAC OSJNB0064P21 genomic sequence,
complete sequence.
AC073166
AC073166.7 GI:12039314
HTG.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 142114)
Buell, C.R., Yuan, Q., Mofatt, K.S., Hill, J.N., Burr, P.C., Hsiao, J.,
Ziemann, V., Pal, G., Bowman, C.L., Fujii, C.Y., VanAken, S.E.,
Bowman, C.L., Craven, B., Utterback, T.R., Khalak, H., Feldblum, T.V.,
Quackenbush, J., White, O., Salzberg, S.L., and Fraser, C.M.
Oryza sativa chromosome 10 BAC OSJNB0064P21 genomic sequence
2 (bases 1 to 142114)
Buell, R.
Direct Submission
Submitted (09-JUN-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
3 (bases 1 to 142114)
Buell, R.
Direct Submission
Submitted (05-JAN-2001) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, rbuelli@tigr.org
4 (bases 1 to 142114)
Buell, R.
Direct Submission
Submitted (28-AUG-2001) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, rbuelli@tigr.org
On Jan 5, 2001 this sequence version replaced gi:11968438.
Address all correspondence to:rice@tigr.org

BAC clone OSJNB0064P21 is from Oryza sativa chromosome 10
The orientation of the sequence is from SP6 to T7 end of the BAC
clone.
Genes were identified by a combination of several methods: Gene
prediction programs including GENSCAN and Genescan (Chris Burge,
http://CCR-081.mit.edu/GENSCAN.html), GeneMark (Mark Borodovsky,
http://genemark.biology.gatech.edu/GeneMark/), and GeneSplicer
(Mihaela Pertea and Steven Salzberg, contact mpertea@tigr.org),
searches of the complete sequence against a peptide database and
the plant EST database at TIGR (http://www.tigr.org/cdb/shm).
Annotated genes are named to indicate the level of evidence for
their annotation. Genes with similarity to other proteins are named
after the database hits. Genes without significant peptide
similarity but with EST similarity are named as unknown proteins.
Genes without protein or EST similarity, that are predicted by more
than two gene prediction programs over most of their length are
annotated as hypothetical proteins. Genes encoding tRNAs are
predicted by tRNAscan-SE (Sean Eddy,
http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are
identified by RepeatMasker (Arjan Smit,
http://ftp.genome.washington.edu/RM/RepeatMasker.html).
location/Qualifiers
1. 142114
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:39947"

```

FEATURES

SOURCE

```

/chrromosome="10"
/map="R2500"
/clone="OSJNB0064P21.16"
/complement(3008..3658)
/gene="OSJNB0064P21.16"
/notes="predicted by genescan"
/complement(<3008..>3658)
/gene="OSJNB0064P21.16"
/complement(3008..3658)
/gene="OSJNB0064P21.16"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAG46103.1"
/db_xref="GI:12039315"
/translation="MWAGDSQARLVIALRLIDPRAAAAPAPLFRHSNDYATP
ARGISVDYFWAPFESNUTRLIHEDRLAPRTDVLVSGMLHMLHVDAAAYGDLA
SVYDAKLSRSLPVPFPMFWLGLPLVNHMLNTDAKKNHNDITLQAYDLEVGORG
LHQRDGPPLLDVGLSRGCGQCTADGMHYDGVYDAVLIHMLALVIESQORI"
/complement(3574..3596)
/rpt_family="GC_rich"
/rpt_family="GC_rich"
/complement(4269..4297)
/rpt_family="GGGA)n"
/rpt_family="GGGA)n"
5613..5633
/rpt_family="GGGA)n"
/complement(5634..6005)
/gene="OSJNB0064P21.1"
/complement(5634..5945)
/gene="OSJNB0064P21.1"
/codon_start=1
/product="histone H4"
/protein_id="AAG46106.1"
/db_xref="GI:12039318"
/translation="MSGRGKGGKGGKGAKKRKYLRNIGITTKPATRLARGGV
KRISGLIYETRGVLKIFLENYIRDAVTTEARAKRTYAMDVYALKKQGTLYIGFG
G"
/complement(7473..7546)
/rpt_family="CGG)n"
/rpt_family="CGG)n"
7556..7630
/rpt_family="GGAGAA)n"
/complement(8238..8259)
/rpt_family="AT_rich"
/complement(8915..8963)
/rpt_family="AT_rich"
/complement(11055..11097)
/rpt_family="AT_rich"
/complement(11244..11264)
/rpt_family="AT_rich"
14313..18336
/gene="OSJNB0064P21.2"
/notes="identical to calcium dependent protein kinase
GI:435466 (Oryza sativa) (Gene 129 (2), 183 189 (1993));
EST C22538, C22540, AA750452, C99271, C22539 from this
gene"
join(14313..15004,15338..15481,15564..15716,15894..16009,
16646..16810,17112..17333,17903..18336)
/gene="OSJNB0064P21.2"
join(14333..15004,15338..15481,15564..15716,15894..16009,
16646..16810,17112..17333,17903..18025)
/gene="OSJNB0064P21.2"
/codon_start=1
/product="calcium-dependent protein kinase"
/protein_id="AAG46110.1"
/db_xref="GI:12039322"
/translation="MGNSSQNGTYGNNYNSNRFQNDPRASRYVNDQNTEDCSGSSR
ASLAGLRQGLNKSPVLGKTPNVELTTLGELGQGGFKTYLCTELSTGCGYACK
TLKSNLCVSDIEDVREIQTIMHLSGKNYITITQYDEDAVAHIMELGAGGLF
SKIKRGHYSERKAELIKIYIGIETCSHGMHMLDLPENFLDLADDEFSVKAD
FGLSVFRPGQVREYVGSPPYIAPLEVRKRGPEADIMTAGIIVALLTGVPFPAAD

```

```

TOSGIYEKVLDRIDIFKSNRWRPISDSKDLIKKMLCPYSEERLKAHEVLKHPWICDN
GVATNRALDPSVLPRLOFSAMNKLGLIOIIRLSEEEIYGLREMFKADTKNS
VTFGEGLKGLKRVSSVPEKDEINLMLEADDTSTINNEEIIAAYSNLKERKHM
AAFTYFDKSGGFTTVDKLOKACHERNEDTFLBMLILEVDNDGQIDVAFETMMQ
SNNFGLQWQTVESLNLVALREAPYR"
/complement(18675..18732)
/rpt_family="GC_rich"
/complement(120605..20634)
/rpt_family="AT_rich"
22393..24891
/gene="OSJNB0064P21.6"
/notes="predicted by genemark hm"
<22393..>24891
/gene="OSJNB0064P21.6"
22393..24891
/codon_start=1
/product="hypothetical protein"
/protein_id="AAG46113.1"
/db_xref="GI:12039325"
/translation="MAGSGERMASSGDLGEYLDRPDAIHRPAASVAIVRSGGDDP
RIVDGRDRPARSGRRSLSSMRPACGETSVSSRRYGTAAALAEYGGKTRAAIY
EMKPVRLISRIQKRCCLLSVEVAGVGPASMDGLAVTRKAETKGAQVOTMPA
TVRGGCADFVEITLTVRCNIVPAGAGTGKPKLBRPRVSVVAEPVAGVGAAYD
VSLVLDLSKSSSEGRVRFMDTAVTISGKATGSLKLKLFOLMDMDGLCLYQOAT
EKVVDVDSVSPALAAHNNKSPVARTSGPPLSASDAISPMBAVYKQIDRLRLEO
GDPYTSILPRKLADDELSGDLGULEYEVVDKGVETVAVYTHARVDLKSLDSVBO
IEALMALNTGKKSPPSPKRVDDQCLDADEBMTVEFLKLEVDNDKGRKLKQPMPT
PRSEBKKAAAPPVVDLGGPGLTAVQTDGGFLVSMNPFDLPLEKXDAAPPKLAQ
VSRPVLPPSSMAATGFDVLQMAAGADEYDKVARGMDNLGKPEBGVFEFGIA
EAVIGRGRTGESSAARSVRVLRLAAVSDGSEERATGWTFAADPETLEEVIAF
SLOKLEAMAVDALMTQAMADEDAPEFVAPAGATTPFDALVSPDESESGSGRV
TVYVALIQDRPSRREAVGAPVAVVGSARILGAAGNGGKPKRKSILVGVQRLCPV
GGAGGAGGASGAEROKLTAMQMALHAPAPAGGRATPTTTTSSQARQRPDVVW
SLSSRVLAGMWLKVTRNPDVAVGATPAGSGGGGn"
/complement(12376..23788)
/rpt_family="GC_rich"
/complement(124616..24639)
/rpt_family="GC_rich"
24853..24903
/rpt_family="GC_rich"
/complement(26149..26194)
/rpt_family="AT_rich"
/complement(26861..26905)
/rpt_family="AT_rich"
/complement(27502..27537)
/rpt_family="GC_rich"
/complement(27628..27663)
/rpt_family="AT_rich"
/complement(28100..28132)
/rpt_family="AT_rich"
/complement(29892)
/rpt_family="TA)n"
29893..29928
/notes="with dinucleotide (AT) repeat (35 to 42 copies)"
34449..38345
/gene="OSJNB0064P21.3"
/notes="similar to sugar transporter GI:1495273
(Arabidopsis thaliana); EST AA33788, D48171, C72446,
AA749646 from this gene"
join(34449..34603,34798..34894,35286..35852,35977..36929,

```

Alignment Scores:

Pred. No.:	Length:	Matches:
6,016-156	142114	552
Score:	2514.00	75
Percent Similarity:	53.77%	109
Best Local Similarity:	47.34%	431
Query Match:	61.69%	9
DB:	Gaps:	

US-10-051-909-32 (1-800) x AC073166 (1-142114)

```

QY 56 MetSerGlyAlaValLeuValAlaIleSerIleGlyAsnLeuGlnGlyTrp 75
DB 34816 ATGGCGGGCGCGCTGCTGGTGGCCATCGCGGCTCCATCGGACATTGGCTGAGGGCTGG 34875
QY 76 AspAsnAlaThrIleAla----- 81
DB 34876 GATTATGCAACCATTCAGAGTAGTTCGTGGTTAGAGAAAAACGAATCTTTTCTTCT 34935
QY 81 ----- 81
DB 34936 GCCAATGTTGTGTGCCATGTGCTTAATGACGCTAAGGCCGTGTTGTGGCAATTAT 34995
QY 81 ----- 81
DB 34996 TGTGATCTTGTGAGATCATGTGATTTGGTGAATCATCTGCAGTAAAAATAGTTT 35055
QY 81 ----- 81
DB 35056 AGTTGTTCTTCTGTTCATGATATGACAAAAAACAATGATTTTTTTTCAGTA 35115
QY 81 ----- 81
DB 35116 GAATTTGAGTCAATTATGAGTAGAGATTTAACAAATGTACGATCTCATGAATGATCTGT 35175
QY 81 ----- 81
DB 35176 ATTATGGAACCTTGATCTGTCTGCATCGTAAAGTTTCAGAAATTTGATGATACCGATG 35235
QY 82 ----- 82
DB 35236 GTTTGACTTCTTGTGACCCCTGTGGACTTGTGGGATGATCTTATACAGTGGCTACT 35295
QY 85 LTYTILELYSLYGLINPHEGLINLEUGLIMANGIUPROTHRVALLIGLILEULIYVALISE 105
DB 35296 GTACATCAAGAAAGAAATTAACCTTGCAGAGCGAGCCCTTATGCAAGGCTCATGCTGTGC 35355
QY 105 rMetSerLeuIleGlyAlaThrIleValThrThrPheSerGlyProLeuSerAspSerIle 125
DB 35356 CATGTGCTCATTTGGGGCGACATCATCAACGAGTCTCTGAGACGATGGCTGATTTCTTT 35415
QY 125 eGlyArgArgProMetLeuIleLeuSerSerIleLeuYrPhePheSerGlyLeuIleW 145
DB 35416 TGGTGGGGCGCCCATCTGATCGCGTGGCTGCTTCTTGTAGGGGCTAGTAT 35475
QY 145 LLeuTrpSerProAsnValIyrValLeuLeuLeuAlaArgPheValAspGlyPheGlyI 165
DB 35476 GCTTGGGGCGCAAAATGTATGTGTGCTTGGCGAGGCTCATTTGAGGGGTTCGGGAT 35535
QY 165 eGlyLeuAlaValThrLeuValProLeuYrIleSerGluIleAlaProSerGluIleAr 185
DB 35536 CCGTTTGGCTGTCAACGCTTGTACATTTGATCATCTCTGAGACGCTGCCCGACGCACTCG 35595
QY 185 gGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMetPheLeuSerTyCy 205
DB 35596 AGGACGTGCTAAACACGCTCCGACGTTCAGTGGGTCTGAGGGAGATGTTCTTTCATACCTG 35655
QY 205 sMetValPheGlyMetSerLeuSerProSerProAspTPPArgIleMetLeuGlyValle 225
DB 35656 CATGTGATTTGGCATGTCTCCATGACCAAGCAAGATTGAGAGATCATCTTGGCGTTCT 35715
QY 225 uAlaIleProSerLeuPhePheGlyLeuThrIlePheYrLeuProGluSerProAr 245
DB 35716 ATCAATACCATCACTTATATCTTTCATGACCATCTTTTACTTACCTGAATCCGCGAG 35775
QY 245 gTriPheValSerIyeglyArgMetAlaGluAlaIySLyValLeuGlnIyLeuArgI 265
DB 35776 GTGGGCTCGTGAACAAGAAAGATGCTGAGCGCAAGCTGTGTTCGAAGGCTGTGGTGG 35835
QY 265 yLyAspAspValSer----- 270
DB 35836 AAGACAAGATGTTTCAGTGTGTGATGTCTTACCAACATTTTCCCTGTGAGATCTAGA 35895

```

```

QY 270 ----- 270
DB 35896 AAGTGAGAAACTATTATATGCTATACGATATAGATATATATATATGAGGATTTCT 35955
QY 271 ----- 35956 CATGTGAAATTAACCTGTACAGAGAAATAGCGCTTCTGTGAAAGTCTGGGGGTGGGA 36015
DB 35956 CATGTGAAATTAACCTGTACAGAGAAATAGCGCTTCTGTGAAAGTCTGGGGGTGGGA 36015
QY 284 LyAspThrSerIleGlnGluIyrIleIleGlyProAlaIleThrGluAlaIleAspLeuV 304
DB 36016 AAGACAAAAATTTAGAAATPACATATTTGACCTGATGATGACCTTCTGATGAAGGAC 36075
QY 304 alThrAspGlyAspIyeglyGlnIleThrLeuYrGlyProGlnGluIyGlnSerTrpI 324
DB 36076 TGGCTCCA---GATCCAGAGAAAGTCAACTGTATGCTGTGAAGAGGCTTATGTGG 36132
QY 324 LeAlaArgProSerIyeglyProIleMetLeuGlySerValLeuSerLeuAlaSerArg 344
DB 36133 TTGCCCGTCTGTTCCAGCGGCAAGTGCATTTGGAAATGATAGTCTCATCTCTCGTC 36192
QY 344 IeglySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlyS 364
DB 36193 ATGTAGATATGCTCATGAGGATTAAGCCCTTGTGATCTGTGTGCACCCCTTTTGGAA 36252
QY 364 eValIleGlyAsnMetProGlnAlaGlyIySerMetArgSerThrLeuPheProAsp 384
DB 36253 GTGTSCATGAGAAAGATGCTGATGATATGGAAGACATCGGAGCACTTTGTTCTTAAT 36312
QY 384 heGlySerMetPheSerValThrAspGlnIleAlaIyAsnGlnIleTPPAspGluAla 404
DB 36313 TTGGCAGCATGTTTGTGTGGCGGAAACAGCAAGCTAAAGGTGATGGATGGATGCTGAGA 36372
QY 404 sLeuLeuIleAspAspAspIyGluIyIyrIleAspAspGlyAlaGlyIyAspYrGlyAspA 424
DB 36373 GT---CAAGGAGAGGCTAATAATTTAGATCAACATGCTGGGGATGATCAATTGAAGATA 36429
QY 424 sLeuLeuIleSerProLeuLeuSerArgGlnAlaThrGlyAlaGluIyLyAspIle---V 443
DB 36430 GCGTCAAAAGCCCACTTATTTCTGTCTCAAGGACAAAGCGTGGAAAGAAAGAGTCCGTCG 36489
QY 443 alIleIleGlyIleIleArgIySerAlaLeuSerMetArgGlnIleThrLeuLeuGlyIug 463
DB 36490 CACCTCATGAGCAATTAATGCTGCTGTGGCA---AGAATGATGATGCTCAAGCAGCAG 36543
QY 463 LyGlyAspGlyValSerSerThrAspIleGlyIyGlyIyTPPGLLeuAlaIyrIySTPS 483
DB 36544 GCGGGAGGCAAGTAAAGCATGCGCATTTGGGGAGTGGCAAGTGGCTTGAATGAAATGGA 36603
QY 483 eGluIyLeuGluIyGluIyAsnGlyIyArgIyGluIyGlyIyPheIyAspValIyrLeuIleI 503
DB 36604 CTGAGAGAGAAAGCTCCAGATGCGCAAAAGAAAGTGGCTTCCAAAGTATCTTACCTTGCATG 36663
QY 503 InGluIyValProGlySerArgArgGlySerIleValSerLeuProGlyIyGlyAspV 523
DB 36664 AAGAGGGGTGTACAGATGATGCGAGGGCTGTATCTCATCTCATTTGCT---GGAGGTGATG 36720
QY 523 alPheGlnGlySerGluPheValIleAlaIleValSerGlnSerIleAlaLeuPheS 543
DB 36721 TTCTCTCTGTGGTGTAGTTCGTCAAGGAGCTGCTTGTGTAGGCCAACCTGCTCTTACT 36780
QY 543 eTyLeuGlyLeuAlaGluProArgMetSerAspAlaIleMetValIleIleProSerGluValA 563
DB 36781 CTAAAGATTTATGAGAAAGCGCTTGTGGCCCTGTATGTGATCATCTCAAGGAG 36840
QY 563 laAlaIyeglySerArgTriPlyAspLeuPheGluProGlyIyValArgArgAlaLeuLeuV 583
DB 36841 TTGCTTAAAGTTCAAAAATGAGGACAGCTTATTCGAACTGAGTGAAGCATGCTGTGTTG 36900
QY 583 alGlyValGlyIyIleGlnIleLeuGln----- 591
DB 36901 TTGGCATGGGATACAAATCTTGCAACAGGTAAACCTTTTCTGATTTATGTCAATGTAC 36960
QY 591 ----- 591

```

```

Db 36961 GAATGATGCTGATGTTGAACCTAATTTCTAGAGCTCCAGATAAATTAATCTGAT 37020
QY 591 -----
Db 37021 TAATGTAAGTAAATCAAAAGCATCTTCATGAGTGTCTTCTAGGTTCTAGCGAG 37080
QY 591 -----
Db 37081 TCTCTAATTAAGCTCTTACTTACTTCTGCGAAAACATGTTAGCTGTACGTGTTT 37140
QY 591 -----
Db 37141 TCTTGCAATATATTTCCATACAAAGCAATGAGCTAACTGCAATAAATACCTTTG 37200
QY 591 -----
Db 37201 ATGATTAGAAATCAAAATGCCATCAATGTTGTTTAACTAGTAAGTCAACCT 37260
QY 591 -----
Db 37261 TTAACATCTTGCAATATCTATTAAGAAACCATTTGGAGCTGGCTATGTCACATGACC 37320
QY 591 -----
Db 37321 TGATTACGGGTATTCATCTGAAGACAGAAACAAATGTCATGATCTTTATTCCTG 37380
QY 591 -----
Db 37381 AATTAATTTGCTGAATATGTTCTTAGTTCGAGAAATTAATCTTCAGAAATTTTCAAGT 37440
QY 591 -----
Db 37441 GGCATTAATCATGTAATTTCTTCTTACTCCCTCCATCTTCTTCCCTGTTAT 37500
QY 591 -----
Db 37501 ATTTATATCAGTCTATTCATGTTTACCTCAGTATAGCATGATGTAATGTCAT 37560
QY 592 -----
Db 37561 CAATTTTCCATTAATGTCAGTGTGTCATTAAGGAGTCTGACTACACCTCCACAA 37620
QY 606 TleuGlulnAlaGlyValAlaValIleuSerLysPheGlyLeuSerSerLys 625
Db 37621 ATTCTTGACAAAGCTGGTGTGTTCTTCTTGCAAAATTTGAGCTTCTCTCATCT 37680
QY 626 AlaSerIleLeuIleSerSerLeuThrIleuLeuMetLeuProCysIleGlyPheAla 645
Db 37681 GCATCATTTCTTATTAGCGGAGCTGACACACTTGTGATGCTTCCAGCATGTGTTGCT 37740
QY 646 MetLeuMetLeuSerSerLysPheGlyValArg----- 655
Db 37741 ATGAGGCTCATGATGATGTCGGAAGAAAGTTAGACTCATATTTTACATTTTCTTGT 37800
QY 655 -----
Db 37801 GATTACATGATCACTTCTAAGTAAAGTATATGTAACATTAATTCGCGTTGTTCT 37860
QY 656 -----
Db 37861 CATTCAGGTTTCTT-CTCTTGCACAAATCCCTATCTTATAGTAGACATGATCTTGG 37919
QY 673 ValIleSerAsnLeuIleAspLeuGlyThrIleuAlaAlaIleuSerThrValSer 692
Db 37920 ATTCTTGCAATATTTCTGATGAGGAGCATGATGTCATGCTTCCAGTCCACAGTCACT 37979
QY 693 ValIleValIleThrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCys 712
Db 37980 GTCAATCTACTCTGCTTCTTCTTCTGATGAGGCTGAGGCTTATTCAAACATCTCTGT 38039
QY 713 AlaGluIlePheProThrArgValArgGlyLeuCysAlaIleCysAlaPheThrPhe 732

```

```

Db 38040 GCAGAGATTTCCGACACCGCTTGTGGCATCTGCATAGCAATCTGTGCCCTAACATTC 38099
QY 723 TrrIleGlyAspIleIleValThrTyrSerLeuProValMetIleAsnAlaIleGlyLeu 752
Db 38100 TGGATCGGTGATATATGTAATGTAACATACACCTCCCGATGATCTCAACGCAATGGACTG 38159
QY 753 AlaGlyValPheSerIleTyrAlaValAlaCysLeuIleSerPheValPheValPheLeu 772
Db 38160 GCTGAGTGTGTTGGAATCTACGAGTGTGTCATACGAGGCTTCTCTTGTGTTTCAAG 38219
QY 773 LysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGly 792
Db 38220 AAGGTCGCGAGACAAAGGCGATGCTCTTGAAGTCATACCGAGTCTCTGTGCGGA 38279
QY 793 AlaLysGlnAla 796
Db 38280 GCAAAGCAGGCC 38291

RESULT 10
LOCUS AE017116
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 10, section 70 of 77 of the complete sequence.
ACCESSION AE017116.1 GI:31433303
VERSION AE017116.1
KEYWORDS
SOURCE
ORGANISM Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 300957)
The Rice Chromosome 10 Sequencing Consortium
In-depth view of structure, activity, and evolution of rice
chromosome 10
Science 300, 1566-1569 (2003)
2 (bases 1 to 300957)
Buell, C.R., Wang, R.A., McCombie, W.R., Messing, J. and Yuan, Q.
Direct Submission
Submitted (05-MAY-2003) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
This is the pseudomolecule for rice chromosome 10, which was
constructed by resolving discrepancies between overlapping BACs,
trimming the overlap regions, and linking the unique sequences to
form a contiguous sequence. Genes in individual BAC clone were
identified by a combination of several methods: Gene prediction
programs, searches of the complete sequence against a peptide
database, and EST databases. Genes with similarity to other proteins
are named after the database hits. Genes without significant
peptide similarity but with EST similarity are named as unknown
proteins. Genes without protein or EST similarity, that are
predicted by more than two gene prediction programs over most of
their length are annotated as hypothetical proteins. Genes
encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy,
http://genome.wustl.edu/eddy/tRNAscan-SE/).
FEATURES
source
location/Qualifiers
1..300957
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="10"
1..12488
/feature="Chromosome Sequence Derivation: nucleotide sequence
in this region was derived from BAC clone OSUNBa040D23
(GB:AC074196)."
<3310..>4569
/locus_tag="OSUNBa040D23.23"
/notes="similar to chloroplast nucleoid DNA binding protein
CND41 GB:BA22813 GI:2541876 (Nicotiana tabacum)"
<3310..>4569
/locus_tag="OSUNBa040D23.23"
misc_feature
gene
mRNA

```


CDS
3310..4569
/locus_tag="OSJNBa0040D23.23"
/codon_start=1
/product="putative nucleoid DNA binding protein"
/protein_id="AAP54833.1"
/db_xref="GI:31433304"
/translation="MGRLLAPMQLVLCISVTTCAAHGRLRGDLDOGNRGRILADAT
AAPGGAVVPLHMSGAHVANFTIGTPQAVSGIVLISGELVWTOCAACRSGCCPKOE
LHPDPSASNTYRAEQCGSPDKSIPTRNCGSGEGVAPSMFGDTFIASTDIAI
GNAEGRALGVAVDGSIDGAMDGSGFVGGRTPWLSVGSNTAFSVCYALHGP
KKSALFLGASATLACGKSNPPTPLIGJASNTSDGSPPTVVOEGIKADYAVAA
ASSGGATVTLQLETRPLSYLDDAYQLLEVVDALGSPSMANPEPEFDICFNAA
VSGVDLVTFFQGAATLTAQPSKYLIGDNGNGVCLSTLSSTRLSDADDGVSIIIGSL
LQENVHFLDEKETSFEPADCCSLT"
/locus_tag="OSJNBa0040D23.24"
/note="predicted by fgenesh"
/join(<5409..5547,5811..5592)
/locus_tag="OSJNBa0040D23.24"
/join(5409..5547,5811..5592)
/locus_tag="OSJNBa0040D23.24"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAP54834.1"
/db_xref="GI:31433305"
/translation="MTTGNLNSNIRICLTKEPCILKHPTCGAMLRQRTSTKASHCP
TEGRKASKSDLSIAVLNDQNPLNQGOTAPLSTIIAAGSTFVARRRNPPFKTKKN
MYAL"
/locus_tag="OSJNBa0040D23.25"
/note="similar to chloroplast nucleoid DNA binding protein
CND41 GB:BA22813 GI:2541876 (Nicotiana tabacum)"
/locus_tag="OSJNBa0040D23.25"
/join(<8062..9183
/locus_tag="OSJNBa0040D23.25"
/join(8062..9183
/locus_tag="OSJNBa0040D23.25"
/codon_start=1
/product="putative nucleoid DNA binding protein"
/protein_id="AAP54835.1"
/db_xref="GI:31433306"
/translation="MRGLLADAAVAGGCVVPYWSQPLVYMANLTGTPOPSAI
IHLAEFWTQSPCRCKQDPLFNRIASSTYRPECGTALCSSVASTSGDVC
SEVEITPFGDTSIGDTFPAIGTATLISAGCANDSNIKQLGAGVVGIGRTWML
VQGNMATAFSYCLAPHGAAGKSKATLISAGSALAGSKAATPLVNTSDSDSYIHL
EGIKFGDVIAPPPGNSVYLPTIFGVSNFLVDAAFOAIKAVTAVGAAPATPKPFP
DLCFPRKAAAAGANSISLPDYVLTFCGAAALTVPBPKXYMYAGNGCYCIAMMSAML
NLTTELSTLGRHGENIHLFDLDEKETSFEPADCCSLT"
/locus_tag="OSJNBa0025B05.4"
/note="similar to chloroplast nucleoid DNA binding protein
CND41 GB:BA22813 GI:2541876 (Nicotiana tabacum)"
/locus_tag="OSJNBa0025B05.4"
/join(<11356..12573
/locus_tag="OSJNBa0025B05.4"
/join(11356..12573
/locus_tag="OSJNBa0025B05.4"
/codon_start=1
/product="putative chloroplast nucleoid DNA binding
protein"
/protein_id="AAP54836.1"
/db_xref="GI:31433307"
/translation="MGRPAVTLFVLCFISVTACSLSEQATRGRLLAGVDAPPAGGA
VAVPIYLSQGLVANFTIGTPQPSAVVDLISGLVWTOCPQCFEODLFPDP
KSTPRGLPCGSHCESIPESRNCSTSDIYEATPKAGDNGKAGTDPATIGAKET
LGFQGVWMDRELKTIIGPSGIVGRTWMLYQWNTAFSVCYALHGP
KKSALFLGASATLACGKSNPPTPLIGJASNTSDGSPPTVVOEGIKADYAVAA
ASSGGATVTLQLETRPLSYLDDAYQLLEVVDALGSPSMANPEPEFDICFNAA
VSGVDLVTFFQGAATLTAQPSKYLIGDNGNGVCLSTLSSTRLSDADDGVSIIIGSL
LQENVHFLDEKETSFEPADCCSLT"
misc_feature
12489..23141
/note="Chromosome Sequence Derivation: nucleotide sequence
in this region was derived from BAC clone OSJNBa0025B05
(GB:AC096782)."
complement(<16252..517307)

mRNA
/locus_tag="OSJNBa0025B05.3"
/note="predicted by fgenesh"
complement(<16252..517307)
/locus_tag="OSJNBa0025B05.3"
/codon_start=1
/locus_tag="OSJNBa0025B05.3"
/product="hypothetical protein"
/protein_id="AAP54837.1"
/db_xref="GI:31433308"
/translation="MWSAAQAGVAAVCVLPFMGLAGMHLNRKVLFPSCALFVSL
AVGHLSPYLSTPLHLAASFHPPISSSSSSAASSCVPLHRSWADAGESSG
GRAMAWPESLSTGCLRLASRSDASLISNGSWWVAGDSQARLLVLRLLPLPAA
AAAPELPFRSDYRATVPARGIVDFWAPAEENLFLLEDRLAARTPDVVLGS
GLMHMLHTDADRYGDLASVYDAKSLRSPLPPPMFMPLGLPLVNMHLNTDARK
VHMNDTIQADYLEVEQRLQROGGRPLLDVGLSRGCGQCTADGMHIDGVYDA
VHMTMLNLTVESQORI"
complement(<18878..519249)
/locus_tag="OSJNB0064P21.1"
/note="similar to histone H4 (H4C13) GI:168499 (Zea mays),
(Plant Physiol Biochem 25, 235 247 (1987)); EST D23040
from this gene"
complement(<18878..519249)
/locus_tag="OSJNB0064P21.1"
complement(18878..519189)
/locus_tag="OSJNB0064P21.1"
/codon_start=1
/product="histone H4"
/protein_id="AAP54838.1"
/db_xref="GI:31433309"
/translation="MSGRGKGGKGLGKGAQRHKKVLRDNIQGITKPAIRRLARRGV
KRIGLLIYEETRGVLTFLFENVIRDAVYTEHARRKVTAMDVYALKRQRTLYGFG
G"
/locus_tag="OSJNBa0025B05.2"
/note="similar to B regulatory subunit of PP2A
GB:AB58900 GI:2160690 (Arabidopsis thaliana)"
/join(<19395..19473,19781..19910,19992..520598)
/locus_tag="OSJNBa0025B05.2"
/join(19395..19473,19781..19910,19992..520598)
/locus_tag="OSJNBa0025B05.2"
/codon_start=1
/product="putative protein phosphatase 2A subunit"
/protein_id="AAP54839.1"
/db_xref="GI:31433310"
/translation="MTVLVTTGRDARIGRVLYHRMRMGACGPPGCGAGODRRRL
VPAAGCRRPPTRRPRPTCTSTRTLRNVDSFLAGLALAFEDSRERDLKY
YHQLYSKLVYERAFMRIMAAALRLFYDTSPEAEHRCVGLLELTCGSLINFAVP
LKEHRAFMARVLLPLHRTTRVHTYHROLSYCGSIRSRHLPPPLAGOMEGGAHR
RARGDRDPLSAPGLPAASAAARFVCFIHPNRRHFSQKKIDFYICTFN"
23142..55358
/note="Chromosome Sequence Derivation: nucleotide sequence
in this region was derived from BAC clone OSJNB0064P21
(GB:AC073166)."
/locus_tag="OSJNB0064P21.2"
/join(<27557..31580
/locus_tag="OSJNB0064P21.2"
/join(27557..31580
/locus_tag="OSJNB0064P21.2"
/codon_start=1
/product="calcium-dependent protein kinase"
/protein_id="AAP54840.1"
/db_xref="GI:31433311"
/translation="MNSCGNGTYGNNYQNSNRFQNDRAFARYVDGNDTEDCYSGSSR
ASLAGALRGILNLSKSPVLYGKTPVNRVLTGLGELGGGKTYLCTEISTGCGYACK
TILKSNLRCSIEDIEVRRLEIQIMHLSGQKNIVITIKOTYDEQGVHIMELCAGSELF

SKIQRGHYSEKKAELIKIIVGIIETCHSHGWHRDKPENFLLLDADDEFSYKAD

Alignment Scores:

Pred. No.:	1.78e-155	Length:	300957
Score:	2514.00	Matches:	552
Percent Similarity:	53.77%	Conservative:	75
Best Local Similarity:	47.34%	Mismatches:	109
Query Match:	61.69%	Indels:	431
DB:	8	Gaps:	9

US-10-051-909-32 (1-800) x AE017116 (1-300957)

```
QY 56 MetSerGlyAlaValIleuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnIleYTP 75
DB 48060 ATGGCGGGGGCGGCTGCTGGTGCCTTCGCGGCTTCATCGGCACTTGTGCGAGGGCTGG 48119
QY 76 AspaAsnAlaThrIleAla-----
DB 48120 GATMAATGCAACCATTCGAGTAAGTTCGTGTAGAGAAAAAAGCAATCTTTTCTCT 48179
QY 81 -----
DB 48180 GCCATAGTTTGTGTCCATGCTCTTATGACGCTAGGGCCCGGTGTGGCGCAATTAT 48239
QY 81 -----
DB 48240 TGTGATCTTGTGAGATCATGATGAGATTGGTGGAATCATCTGCAAGTAAATAGTTT 48239
QY 81 -----
DB 48300 AGTTGTCTCTCTGCTGATGATATACAAAAAACAATGATTTTTTTTTCAGTA 48359
QY 81 -----
DB 48360 GAATTTTGAATCAATTATGATGAGATTAAACAAATGACGATCTCATGAATGATCTGT 48419
QY 81 -----
DB 48420 ATTATGAACTTGATACGTCTGATGATGAGTTTTCAGAAATTTGATGATACCATG 48479
QY 82 -----
DB 48480 GTTTTGACTTCTTGTGACCCCTGTGGAAGTGTGGGAATGATCTTATCAGGCGGCTACT 48539
QY 85 -----
DB 48540 GTATCTTCAAGAGAAATTCACCTTCAAGAGCGGCCCTTATGAGAGCGGCTGATGCTG 48599
QY 105 MetSerLeuIleGlyAlaThrIleValIleThrPheSerGlyProLeuSerAspSerIle 125
DB 48600 CATGTCGCTCATTTGGGGGAGCATATCATCAACGCTTCTGTGAGAGAGTGGCTGATTTCT 48659
QY 125 eGIYArghrProMetLeuIleLeuSerSerIleLeuYrPhePheSerGlyLeuIleIle 145
DB 48660 TGGTAGGGGGCCCATGCTGATGCTGCGGCTGCTCTTCTTCTTCTTCTTCTTCTTCTTCT 48719
QY 145 tLeuYrPheProAsnValIleValIleLeuLeuLeuAlaArgPheValIleArgPheGlyIle 165
DB 48720 GCTTTGGCGCGCAAAATGTGTATGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 48779
QY 165 eGIYLeuAlaValIleThrLeuValIleProLeuYrIleSerGluIleAlaProSerGluIle 185
DB 48780 CGGTTTGGCTGTCAAGCTTGTGTACATTTGTACATCTCTGAGACTGCCCGCAGGAGATCAG 48839
QY 185 gGIYLeuLeuAsnThrLeuProGlnPheSerGlySerGlyIleMetPheLeuSerIleYrCY 205
DB 48840 AGGATGTCTAAACAGCGTCCCGCAGTTCAGTGGCTCTGGAGGAGATGTTCTTCTTCTTCT 48899
QY 205 eMetValIlePheGlyMetSerLeuSerProSerProAspTrpArgIleMetLeuGlyValIle 225
DB 48900 CATGGATTTTGGCGAGATGCTCTTATGCAAGCAATTTGAGAGATCATGCTTGGGCTTCT 48959
QY 225 uAlaIleProSerLeuPhePheGlyLeuThrIlePheYrLeuProGluSerProXr 245
```

```
DB 48960 ATCAATACCATCATATTATATCTTTGATGATGCAATCTTTTACTTACTGATGATGCGGAG 49019
QY 245 gTTrleuValSerIleGlyAlaGlyMetAlaGluAlaIleValIleValIleValIleValIle 265
DB 49020 GTGGCTGTGTGAGCAAAAGAGAAATGGCTGAGGCCAAGGCTGTGTGCAAGGCTTGGCTGG 49079
QY 265 YLysAspAspValSer-----
DB 49080 AAGAGAAATGTTTCAGGTTTGTATGTTCTTACCCAGCATTTTCCCTGTGAGATCTTACA 49139
QY 270 -----
DB 49140 AAGTGAGAAATATTATATGCTATACATATAGTATATATATATGATGATGATGATTTCT 49199
QY 271 -----
DB 49200 CATGTTGAATTAACCTGTGACAGAGAAATGCGCTTCTGCTGAAAGTCTGGGGGTTGGGA 49259
QY 284 LYAspThrSerIleGluGluIleIleIleGlyProAlaThrGluAlaAlaAspAspLeuVal 304
DB 49260 AAGACACAAAAATGAGAAATCATATTTGACCTGATGATGAGCTTCTGATGAGAGGC 49319
QY 304 aThrAspGlyAspIleGluGluIleIleThrLeuYrGlyProGluGluIleGlnSerIleTPI 324
DB 49320 TGGCTCCA---GATCAGAGAAATCAATCAATGCTATGCTCTGAGAGAGAGCTTATGCTGG 49376
QY 324 LeuAlaArgProSerIleGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgH 344
DB 49377 TTTGCCGCTCTGTTTCAACGGGCAAGTGCATGGAAAGTGTATGCTCATCTCTCTGTC 49436
QY 344 IAspIleSerMetValAsnGlnSerValIleProIleMetAspProIleValIleThrLeuPheGlyS 364
DB 49437 ATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 49496
QY 364 eValIleIleGluAsnMetProGlnIleGluIleGlySerMetArgSerThrLeuPheProAsp 384
DB 49497 GTGTCAATGAGAAATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 49556
QY 384 heGlySerMetPheSerValIleThrAspGlnIleAlaIleValAsnGluIleIleIleIle 404
DB 49557 TTTGGCAGCATGTTTATGTTGCGGAGAACAGCAAGCAAGTAAAGTATGATGATGATGATGAT 49616
QY 404 eIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 424
DB 49617 GT---CAACGGAGAGGTGAAATATGATGATGATGATGATGATGATGATGATGATGATGAT 49673
QY 424 eIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 443
DB 49674 GCTTCCAAAGCCCACTTATTTCTGCTCAAGGAGCAAGGCTGAGAGAGAGAGATGCTGCTG 49733
QY 443 aLHsHsIleGlyIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 463
DB 49734 CACCTCATGAGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 49787
QY 463 IYGIYAspGlyValIleSerSerThrAspIleIleGlyIleIleIleIleIleIleIleIleIle 483
DB 49788 CGGGGAGGAGCAATGAGCAATGAGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 49847
QY 483 eGIYLeuGluIleGluIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 503
DB 49848 CTGAGAGAGAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 49907
QY 503 IYGIYValIleProGlySerArgIleIleIleIleIleIleIleIleIleIleIleIleIleIle 523
DB 49908 AAGAGGCTGTGACAGTGTATGACAGGCTCTTATATCTTATCTTATCTTATCTTATCTTAT 49964
QY 523 aPheGluIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 543
DB 49965 TTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 50024
QY 543 eTLYGlyLeuAlaGluProArgMetSerAspAlaIleMetValIleIleProSerGluValA 563
```

Db 50025 CTAAAGAAATGATGAGCAACGCTTGCCCTGCTATGGTGCATCCATCCAGGAG 50084
 QY 563 laAlaIysGlySerArgTplysAspLeuPheGluProGlyValAlaGSAAlaLeuLeuV 583
 Db 50085 TTGGTAAAGATCCAAATAGGAGGAGACTTATGCAACCTGAGTGAACATGCTGCTGTTG 50144
 QY 583 aAlGlyValGlyIleGlnIleLeuGln----- 591
 Db 50145 TTGGCATAGGATACAAATCTCGCAACAGSTAACCTTTTCTGATTTATGTCATGAC 50204
 QY 591 ----- 591
 Db 50205 GAATGATGCTAGTGTGAACTTAATTTCTAGTACGCTCAGATTAATTAATCTCTGAT 50264
 QY 591 ----- 591
 Db 50265 TAAATGTAAGTGTAAATCAAAAGCATCTTCATGAGTCTCTTCTAGGTTCTTAGCGAG 50324
 QY 591 ----- 591
 Db 50325 TCTTCATTAAGCTCTTACTTACTTACTTCTGCAAAAACATGTTAGCTGTACGTTT 50384
 QY 591 ----- 591
 Db 50385 TCTTCATTAATATTCATCAACAGCAAGCATGAGTAACTGCAATAATACACCTTGG 50444
 QY 591 ----- 591
 Db 50445 ATGATTAAGATACAAATGCCATCAATGTTGTTTACTTTTACACTAGTAAGTCAACCT 50504
 QY 591 ----- 591
 Db 50505 TAAATCTTGAGATATTCATTAATTAAGAAAACATTGGAGCTGGCTATGTCATGAC 50564
 QY 591 ----- 591
 Db 50565 TGATTAACGGTATTCATTCAGAACAGCAAAACAAATGTCATGATCTTTTATTCCTG 50624
 QY 591 ----- 591
 Db 50625 AATTTATTTGCTGATATGTTCTTACTTTCGCAAAATACTTCAGAAACTTTTACAGT 50684
 QY 591 ----- 591
 Db 50685 GGCATTAATCATGATATTCCTTTTACTTCCCTCCATCTTCTATTTTCCCTTGTAT 50744
 QY 591 ----- 591
 Db 50745 ATTATTAATCACTATTCATGATGTTTAACTCACTATGATGATGATGATGATGAT 50804
 QY 592 ----- 592
 Db 50805 CAATCTTTCCATTAATGTCAGATTTGCTGCAATTAAGGATTCCTGACTACACTCCACAA 50864
 QY 606 IleLeuGlnGlnAlaGlyValAlaValIleLeuSerIysPheGlyLeuSerSerAlaSer 625
 Db 50865 ATCTTTGAGCAACGCTGGTGTGTTCTTTCTTTCGCAAAATGAGCTTACCTCCATCT 50924
 QY 626 AlaSerIleLeuIleSerSerLeuThrIleLeuMetLeuProCysIleGlyPheAla 645
 Db 50925 GCATCTATTTCTTATTAAGCGCACTGCAACCTTGCTGATGCTTCCAGCATGTGTAATCT 50984
 QY 646 MetLeuLeuMetAspLeuSerGlyArgArg----- 655
 Db 50985 ATGAGGCTCATGATATGCTGGAAGAGTTAGACTTCATATTTTACATTTTCTTCTGT 51044
 QY 655 ----- 655
 Db 51045 GATTAATCATGATCTTAAATGAAGATTAATGCTAACATTAATTCGCTGTTTCT 51104
 QY 656 -----PheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeu 672
 Db 51105 CATTCAGGTTCTT-CTCCTTGCAACAACTCCATCTGATAGTAGCACTAGCTATCTTG 51163

QY 673 ValIleSerAsnLeuIleAspLeuGlyThrIleuAlaHisAlaLeuLeuSerThrValSer 692
 Db 51164 ATTCGTGCATATTCCTTGATGAGGAGCAATGGTTACATGCTCAGCTCCACAGTCACT 51223
 QY 693 ValIleValIlePheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCys 712
 Db 51224 GTCACTACTCTACTCTCTCTTTGTATGATGGGTTGGGCTATTTCCAAACATTTCTGT 51283
 QY 713 AlaGluIlePheProThrArgValIleGlyLeuGlyValIleAlaIleCysAlaPheThrPhe 732
 Db 51284 GCAGAGATTTTCCGACCAACCGCTGCGCATCTGCAATGCAATCTGTGCTTCAACATTC 51343
 QY 723 TrpIleGlyAspIleIleValIleThrTyrosLeuProValMetLeuAsnAlaIleGlyLeu 752
 Db 51344 TGATCGGTATATCATTTGTATCATACATCACTCCCTCCGATGCTCCACGCAATGAGCTTC 51403
 QY 753 AlaGlyValPheSerIleTyraIleValIleCysLeuIleSerPheValPheValPheLeu 772
 Db 51404 GTGGAGTGTTTGGATCTACGAGTGTGTGCAATGCTGCTTCTCTGTTCTTCTCATG 51463
 QY 773 LysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGly 792
 Db 51464 AAGGTCCGAGACAAAGGCGATGCTCTTGAACTCATCAACCGAGTTCTTCTGTGCGGA 51523
 QY 793 AlaIysGlnAla 796
 Db 51524 GCAAGCAGGCC 51535
 RESULT 11
 ATRF23E12
 LOCUS ATRF23E12
 DEFINITION Arabidopsis thaliana DNA chromosome 4, BAC clone F23E12 (SSA project).
 ACCESSION AL022604
 VERSION AL022604.1 GI:3080406
 KEYWORDS
 SOURCE
 ORGANISM Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1
 AUTHORS Bevan, M., Hilbert, H., Braun, M., Holzer, E., Brandt, A., Duesterhoeft, A., Hohnsels, J., Mewes, H.W., Mayer, K.F.X. and Schueller, C.
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 86710)
 AUTHORS EU Arabidopsis sequencing project.
 TITLE Direct Submission
 JOURNAL Submitted (01-APR-1999) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: schueller@mpi.biochem.mpg.de, mayer@mpi.biochem.mpg.de, Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk
 COMMENT Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>.
 FEATURES
 source
 1..86710
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /variety="Columbia"
 /db_xref="taxon:3702"
 /chromosome="4"
 173..1890
 /gene="F23E12.10"
 join(4173..374,456..919,1089..1610,1678..1890)
 /gene="F23E12.10"
 /codon_start=1
 /product="hypothetical protein (fragment)"
 /protein_id="CAA18726.1"

```

/db_xref="GI:3080407"
/db_xref="SPTREMBL:O65486"
/translation="FNAVAGCFPAVESEVLWPSSTSDSTSLPTTSMVLISAVTASY
VSEIGLSEPAKVDYDFRSRGGELPLYLALCGVSLASRQTSMTGAVSLINK
DAGI PKAVFPVPMGGLSVGIIALVPEUTLXGROVDILLEKRPYKGSADILLQVYA
VKIATATMCRAAGLVGYTAPSLFI GAGAGAYGKFTGLALQNDPFLSTLEVASPO
AYGLVGMATLAVCOVPLTAVLLFELTDYRI VLPPLGAVGMSWITSQSRQET
RETRERKRKQSEAVQSLTSSDBESTNNLCEVESLCLDSDLNQSEBIPKSI FVSEA
MRTRPATVMTSLEALTRMLIEKOSCALIVDPNIPIGLITLSDIOEFKARKEGN
NRPKDIFVNDICSRSGCKVPMWTPTPMDLLAQTIINNKHELSHVAVVSIDAPRI
HPGVLDRECIITLRR"
exon
173..374
/gene="F23B12.10"
/intron
375..455
/gene="F23B12.10"
/exon
456..919
/gene="F23B12.10"
/intron
920..1088
/gene="F23B12.10"
/exon
1089..1610
/gene="F23B12.10"
/intron
1611..1677
/gene="F23B12.10"
/exon
1678..1890
/gene="F23B12.10"
/intron
1891..4428
/gene="F23B12.20"
join(3124..3241,3323..3465,3595..3696,3945..4107,
4211..4428)
/gene="F23B12.20"
/translation="similarity to dihydroflavonol-4-reductase, Gerbera
x sp., PIR2:S35189"
/codon_start=1
/product="putative protein"
/protein_id="CA18727.1"
/db_xref="GI:3080408"
/db_xref="SPTREMBL:O65487"
/translation="WDOAKGVQVYAGSGFLASMLYKLLLEGYEVIGTVDPNGNEK
LAHLMKEGAKERLRILVAKDLMEBSFTNATMGCGVHTHSPEILRPALICTLNVLR
SCRKPSLKRVLITSSSTVWYALSKTLAEQAWKFESENGIDLTVPVPSLVGSLP
PDLCSITADIVGLKGETEKQWGMQGVYHIDVARTHIIVFEHAAQGRYICSSNV
ISLEIVSFLSARVPSLPIPKRFTDPLN"
exon
3124..3241
/gene="F23B12.20"
/intron
3242..3322
/gene="F23B12.20"
/exon
3323..3465
/gene="F23B12.20"
/intron
3466..3594
/gene="F23B12.20"
/exon
3595..3696
/gene="F23B12.20"
/intron
3697..3944
/gene="F23B12.20"
/exon
3945..4107
/gene="F23B12.20"
/intron
4108..4210
/gene="F23B12.20"
/exon
4211..4428
/gene="F23B12.20"
/intron
4429..5013
/gene="F23B12.30"
/complement(4939..5013,5109..5195,5307..5342,
5451..5651,5725..5811,6161..6163))
/gene="F23B12.30"
/complement(join(4939..5013,5109..5195,5307..5342,
5451..5651,5725..5811,6161..6163))
/gene="F23B12.30"
/translation="Contains Clathrin adaptor complexes small chain
signature [VYKRYASLYF]
contains EST gb:T44219"
/codon_start=1
/product="clathrin assembly protein AP19 homolog"
/protein_id="CA18728.1"
/db_xref="GI:3080409"
/db_xref="SPTREMBL:O23685"
/translation="MIHFVLVSRGQVRLTKWSPYAKERSKVIRESGLVILNRGP
KLGNFVEMRGYKVVYKRYASLYFCMCIDQDNELEVEIITHYVEIILDRFGVCELD
LIFPHKAYIILDELILAGELQESSKKTVAIRIISADQLYEVAAKEBASISNITQAOT
NR"
exon
5014..5108
/gene="F23B12.30"
/intron
5109..5195
/gene="F23B12.30"
/exon
5196..5306
/gene="F23B12.30"
/intron
5307..5342
/gene="F23B12.30"
/exon
5343..5450
/gene="F23B12.30"
/intron
5451..5651
/gene="F23B12.30"
/exon
5652..5724
/gene="F23B12.30"
/intron
5725..5811
/gene="F23B12.30"
/exon
5812..6160
/gene="F23B12.30"
/intron
6161..6163
/gene="F23B12.30"
/exon
6164..7352
/gene="F23B12.40"
/complement(7352..7582)
/gene="F23B12.40"
/complement(7352..7582)
/gene="F23B12.40"
/codon_start=1
/product="hypothetical protein"
/protein_id="CA18729.1"
/db_xref="GI:3080410"
/db_xref="SPTREMBL:O65488"
/translation="MPLVRKNGVYMRIGSPWLHSENLVGDGVAPPRGRDEIGEEV
GGETNVVSESCGSEGNQDAKTTPEGBADRI"
exon
7353..7582
/gene="F23B12.40"
/intron
7583..8233
/gene="F23B12.40"

```

gene /gene="F23B12.50"
complement (8233..9045)
/gene="F23B12.50"
complement (8233..9045)
CDS /note="similarity to DNA-binding protein PDI, Pisum
sativum"
/codon_start=1
/product="putative protein"
/protein_id="CA18730.1"
/db_xref="GI:3080411"
/db_xref="SPTREMBL:O65489"

Alignment Scores:

Pred. No.:	1,866-155	Length:	86710
Score:	2502.00	Matches:	523
Percent Similarity:	68.91%	Conservative:	91
Best Local Similarity:	58.70%	Mismatches:	124
Query Match:	61.40%	Indels:	153
DB:	8	Gaps:	12

US-10-051-909-32 (1-800) x ATF23B12 (1-86710)

```

QY 52 LeuGluAspLys---MetsSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsn 70
DB 39649 ATAGAAGACGAATATATGAGTGTGAGCTGTGTTGCTATTTGCTGCTGCTGTTGGCAAC 39708
QY 71 LeuLeuGlnGlyTTPAspAsnAlaThrIle----- 80
DB 39709 TTGTATCAAGATGGATGATACCAACTATTCAGGTATGATATCTTGTATCATCA 39768
QY 80 ----- 80
DB 39769 ATTCGTTATGAAATTATAGATCCATTAAGTCAGGATTTTATGATACATTACACGA 39828
QY 81 -----A1 81
DB 39829 TCTTTAGCTTATCTCTTTGCTGTGATGATGATTTTAAGTTTCTTTGGTACTCTGC 39888
QY 81 AAlaAlaValLeuTyrIleValYsGluPheGlnLeuGlnAsnGluProThrValGluG1 101
DB 39889 AGGACCTGTGTGATCATATAAAGAGATTTATTTGAGAGTATTCATCCATCAGTGAAG 39948
QY 101 YLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrThrPheSerGlyProLe 121
DB 39949 TCTATTTGTGGAGTGTCACTTATGTGTCCTGATCTGATTAACAACGCTCTGGAGGGGT 40008
QY 121 userSerSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSe 141
DB 40009 AGCTATTTGGCTGTGTCGCCGCCATGCTATATTTGCTCAATTTCTTACTTTGTTGG 40068
QY 141 rGlyLeuIleMetLeuTTPSerProAsnValTyrValLeuLeuLeuAlaArgPheValAs 161
DB 40069 TTCTCTAGTAATGCTATGCTCTCCCAATGTTATGTTGCTCTTACGAGAGCTTTAGA 40128
QY 161 pGlyPheGlyIleGlyLeuAlaValThrIleValProLeuTyrIleSerGluIleAlaPr 181
DB 40129 TGGATTTGGGTTGGCTGTGTGTCACACTTGTCTTATTAATATCTGAACCTGCACC 40188
QY 181 oSerIleuIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyIleMetPh 201
DB 40189 ACCTAGATAGGAGGACTGTTAATACGCTACCGACGTTCACTGCTCTGGAGGATGTT 40248
QY 201 eLeuSerTyrCysMetValPheGlyMetSerLeuSerProSerProAspTPArgIleMe 221
DB 40249 CTTATCTTAAGTATGCTGTTCGGAATGCTGTGATGCTCATACCTAGCTGAGATTTAT 40308
QY 221 tLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeuPr 241
DB 40309 GCTTGGTGTCTTTTCATCTCCCTGCTTTTCTTCTTCTCAGAGGCTTCTTCTTGC 40368
QY 241 oGluSerProArgTTPLeuValSerIleYsGlyArgMetAlaGluAlaIleValLeuG1 261

```

```

DB 40369 CGAGTCCCAAGGTGCTCGTGAACAAAGTCGATGCTTGAACAAAGCGGTTCTTCA 40428
QY 261 nLyLeuAlaArgGlyYsAspAspValSer----- 270
DB 40429 GAGACTGTGTGTGCGCAAGATGTGTCTGTACTTAAATCGTTCTGACTTAATTTCT 40488
QY 271 -----GlyGluLeuS 274
DB 40489 TTCATGAAACCAATCTTTCTATTGAATGCTTATGTGCAATGATATAGTATGATG 40548
QY 274 eLeuLeuLeuGlnGlyLeuGlnValGlyIleAspThrSerIleGlnGluTyrIleIleG 294
DB 40549 CTTTGTGTGTGAGGCTTCTGCAATGAGGTGAACACCATATGAGAAATATATATTG 40608
QY 294 lYProAlaThrGluAlaIleAspAspLeuValThrAspGlyAspGlyGlnGlnIleThrL 314
DB 40609 GTCCCGCGAGTAAAGTATCTGATATCATGATATAGCTGTGATAGATCAAAATTAAGT 40668
QY 314 euryGlyProGlnGluGlyGlnSerTyrIleAlaArgProSerIleGlyProIleMetL 334
DB 40669 TATATGTGCAAGAAAGAGCTGAGTGTGCTTACGCCAGTCAAGCA----- 40718
QY 334 euGlySerValLeuSerLeuAlaSerArgHisGlySer---MetValAsnGlnSerValP 353
DB 40719 --GGAAGACATATAGTGTGTTTGTCTCGCCATGGAAGTACATGACGAGCAAGGCT 40776
QY 353 rOleuMetAspProIleValThrIleuPheGlySerValHisGluAsnMetProGlnAlaG 373
DB 40777 CATGATGATGATCTCTGTGTCACTGTGTTGGAGCTGTCAAGAAAGATGCCGACACT- 40835
QY 373 lYGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspG 393
DB 40836 --GGAAGATAGAGAGTCCCTGTGTTCCACACTTTTGGAGATGTTTCAGTGTTGAGGGA 40893
QY 393 lnhIleAlaYsAsnGlnGlnTTPAspGlnGluAsnLeuHisAlaYsAspAspGlnGluTyrA 413
DB 40894 ATCAACCAAGACATGAAGATGGATGAAGAGATCTTGTGAGAGTACGATTAATC 40953
QY 413 lAserAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgG 433
DB 40954 CATCCGAC---CARGAGATATTTCTGAAGATGATCTTCACTTCCGTATCTCAGTGC 41010
QY 433 lnhIleThrGlyAlaGlnGlyYsAspIleValHisIleGlyHisArgGlySerAlaLeuS 453
DB 41011 AAAGCAGACAGATGAG---AAAGACATGCTCTCACATGCTCAGAACTTTCTACT 41067
QY 453 eMetArgArgGlnThrLeuLeuGlnGlyGlyAspGlyValSerSerThrAspIleG 473
DB 41068 TCAGACATGGAAGTCAAGTGCAGGAGACTCAAGGGAAGGCGGATGATGGGATTTG 41127
QY 473 lYgIYgIYTPGlnLeuAlaThrLysTTPSerGlyLysGlnGlyGluAsnGlyArgLysG 493
DB 41128 GAGGTGATGAGCAAGTGCAGTGAATGAGCAAGAAAGATGATCGGACAGAAAG 41187
QY 493 lGlyGlyPheLysArgValTyrLeuHisGlnGlnGlyValProGlySerArgArgLys 513
DB 41188 AAGGTGGTTTAAACGATATATCTTCATCAAGAGTTTCCAGCATCTGAGTGGCT 41247
QY 513 eTlleValSerLeuProGlyGlyGlyAspValPheGlnGlySerGluPheValHisAla 533
DB 41248 CAATTTGTTCAATCTCGTGTGTGATGAACCGGTGAG---GCAGATTTGTATCAACGCT 41304
QY 533 lAlaIleValSerGlnSerAlaLeuPheSerIleYsGlyLeuAlaGluProArgMetSer 553
DB 41305 CTGCTTTGGTTAGCAACAGCTCTTAATTCAAAGACTTTCACAAAGACATCAATG 41364
QY 553 sPaIleAlaMetValHisProSerGluValAlaIleYsGlySerArgTTPAspLeuP 573
DB 41365 GTCTGCTATATGATCAATCCGAA---ACAATTAAGGTCATTTGGATATATCTTC 41421
QY 573 heGluProGlyValArgArgAlaLeuLeuValGlyIleGlnIleLeuGln--- 591
DB 41422 ATGATCCTGAGAGTCAACGCTGATTAATGATGAGAGTTGAGCTCAATATCTTACAGCAG 41481

```

QY	591	-----	591
Db	41482	TAGCTGATAGCCACATCAAAATTAAGTGAATACAGAGAAATAATGCGTTTATGCTA	41541
QY	592	-----	-ginphealagllyleasnglyValLeuT
Db	41542	ATTCTATGTTTCTCGCTTCCTCTTGGAAMVAGCGTTCTCAGGCATCAACGAGGTTCTTT	41601
QY	601	YTYTYTHrPROGInIlleuGlUGInAlaaglyValAlaValIlleuSerLysPheGlyL	621
Db	41602	ACTACACACCGCAATCTTGAAGAGCGGGGTGCGGAATCCATACATGCAATGCGGGA	41661
QY	621	euserSerAlaSerAlaSerIlleuIlleuSerSerLeuThrThrLeuMetLeuProc	641
Db	41662	TTAGTCTTCTCGTCGACATCTCTACTTAAAGTGATTAAGACACTTGTGATGTAACTCG	41722
QY	641	ysIlleGlypheAlaMetLeuLeuMetAspLeuSerGlyArg-----	654
Db	41722	CAATAGCGTGTGACATAGAGCTCATGATCTTTCTGTGTCGAAGGTATTAACCTTAAT	41781
QY	654	-----	654
Db	41782	TYCCTTGTCTTACATACATATAACGAACAGCTAAATGCGTTATGAATATCTT	41841
QY	655	-----	-ArgPheLeuLeuLeuGlyThrIlleProIlleuIlleAlaSerLeu
Db	41842	TTATTTATATTTGACAGACCTTGCTTCTACACACATACCAATCTGTATGATCTCTA	41901
QY	670	ValIlleuValValSerAsnLeuIlleAspLeuGlyThrLeuAlaAlaLeuLeuSer	689
Db	41902	TTGGTTTATGATATCTCAATCTGTTCATCATGAACAGCATGTGACCGGGGCTTATCA	41961
QY	690	ThrAlaSerValIlleValTyrPheCysCysPheValMetClyPheGlyProIlleProAn	709
Db	41962	ACCGTAAGCGGTGGCTCACTCTGCTTCTGTGTGATGGGTTTGCGTCTGCTCCAAAC	42021
QY	710	IlleuCysAlaGluIllePheProThrArgValArgGlyLeuCysIlleAlaIlleCysAla	729
Db	42022	ATCTCTGTTCCAGAGATTTTTCACATCCGAGTCGCGGATCTGCATCCGATCTCGCA	42081
QY	730	PheThrPheThrIlleGlyAspIlleIlleValThrTyrSerLeuProValMetLeuAla	749
Db	42082	CTCACCTTTCGATCTGTGACATATATGTCATTAACGTCTCCCGGCTCAATATCC	42141
QY	750	IlleGlyLeuAlaGlyValAlaPheSerIlleTyrAlaValAlaCysLeuIlleSerPheValPhe	769
Db	42142	ATTGAGACTAGCTGGATGGTGTGGAATGATGACGCAATCGTAAATGTGCATTTCATGAGGTCCTT	42201
QY	770	ValPheLeuLysValProGluThrLysGlyPheProLeuGluValIlleThrGluPhePhe	789
Db	42202	GTCGTTCAATTAAGATCCCGGAAACTTAAGGCAATGCACTTGAATGATCATCAGAGTCTTT	42261
QY	790	AlaValAlaGlyAlaLysGlnAlaAlaAla	798
Db	42262	TCTGTGAGCTAGACAAAGCTGAGCT	42288
RESULT 12			
ATCIRIV83/c			
LOCUS	ATCIRIV83	197859 bp	DNA linear
DEFINITION	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 83.		
ACCESSION	AL161587		
VERSION	AL161587.2	GI:7270470	
KEYWORDS			
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
AUTHORS	Murphy, G., Ridley, P., Hudson, S., Mewes, H.W., Lemcke, K. and Mayer, K.F.X.		
JOURNAL	Unpublished		

	REFERENCE
JOURNAL REFERENCE AUTHORS	2 (bases 16641 to 103350) Hilbert,H., Braun,M., Holzer,E., Brandt,A., Duesterhoeft,A., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.
JOURNAL REFERENCE AUTHORS	Unpublished 3 (bases 98352 to 125759) Robben,J., Gylmonpez,B., Voickaert,G., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.
JOURNAL REFERENCE AUTHORS	Unpublished 4 (bases 120761 to 197859) Rose,M., Hempel,S., Entlian,K.-D., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.
JOURNAL REFERENCE TITLE JOURNAL	5 (bases 1 to 197859) EU Arabidopsis sequencing project. Direct Submission Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG; E-mail: lemcke@mips.biochem.mpg.de,mayetmips.biochem.mpg.de,project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bsrc.ac.uk
COMMENT	Information on performance analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/ this fragment has an overlap with ATCHRIVe2 at the 5' end and an overlap with ATCHRIV84 at the 3' end. Location/Qualifiers
FEATURES source	1..197859 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /variety="Columbia" /db_xref="taxon:3702" /chromosome="4" 11339..12705 /gene="ATg35190" join(11339..11494,11536..11664,11853..11952,12119..12189, 12314..12457,12499..12705) /note="similarity to ctf, Mycobacterium leprae, gb:U15180- contains EST gb:T45691, T22640, T75954" /codon_start=1 /product="putative protein" /protein_id="CAAB0236.1" /db_xref="GI:7270471"
gene CDS	/translation="MEIKSRFKVCVFCSSSGKRCYDAATDLAQLVLCLINL ESLENLKVTRRLNLYVGGSIGMLGVSOAHVEAGGHVGLAYQIYDLFTLTETYG EVILVAIMRDKAMARMASDPFIALPGSYGLEELVIWAOLGHPKGALNVDG YYNIALTFRIDKAVDGRIRKSQRHLTPISAPRAKELVOKLETLMKNNDKRSDAILL PFDDIPQAKPVPNDGVIAKRWVKTKQQPOOQOOVVFCSTNTQTETAL"
exon intron exon intron exon intron exon intron exon intron	11339..11494 /gene="ATg35190" /number=1 11495..11535 /gene="ATg35190" /number=1 11536..11664 /gene="ATg35190" /number=2 11665..11852 /gene="ATg35190" /number=2 11853..11952 /gene="ATg35190" /number=2 11953..12118 /gene="ATg35190" /number=3 12119..12189 /gene="ATg35190" /number=4 12190..12313 /gene="ATg35190" /number=4


```

Db 63402 TCATATGTCGCGATGTCATTAATGTCCTACCTGATTAACAATGCTCTGAGGGGT 63343
Qy 121 userAspserIleGlyArgArgProMetLeuIleuSerSerIleuTyrrphePhe 141
Db 63342 AGCTGATGGCTGGTGGCCGCGCCATCATATATGTCCTCAATCTCTCATCTTGTGG 63283
Qy 141 rGlyLeuIleMetLeuTyrSerProAsnValTyrrValLeuLeuAlaArgPheValAs 161
Db 63282 TTCTCTAATATGATATGTCCTCCGAATGTTATGTGTCTTGAAGAAGTCTTGA 63223
Qy 161 pGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrrIleSerGluIleAlaPr 181
Db 63222 TGGATTTGGGGTGGTCTTGTGTCACACTTGTCTTATTAATATCTGAGACTGCACC 63163
Qy 181 oSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMetPh 201
Db 63162 ACTGAGATATAGGGAGCTGTGTAATACCTACCGAGTTCACCTGCTCGAGGGATGTT 63103
Qy 201 eLeuSerTyrrCysMetValPheGlyMetSerLeuSerProSerProAspTyrArgIleMe 221
Db 63102 CTATATCTACTGATATGTCCTTTCGGAATGTCGTATGTCATCACTGATGAGATTCAT 63043
Qy 221 tLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrrLeuPr 241
Db 63042 GCTTGGTGTCTTTCATCCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCC 62983
Qy 241 oGlySerProArgTyrPheValSerLeuGlyArgMetAlaGluAlaValysValLeuG 261
Db 62982 CGAGTCCCAAGGTCGTCGTCGAGCAAGTCGTAAGTCCTTGAAGCAAGGGTCTTCA 62923
Qy 261 nLyseLeuArgGlyLyAspAspValSer-----GlyIleLeuS 270
Db 62922 GAGACTGCTGTGTCCGCAAGATGTCCTGATTAATCTTCTGACTCTTATTCCT 62863
Qy 271 -----GlyIleLeuS 274
Db 62862 TTCAATGAACCAATCTTTCATATGATGTCCTATGTCGCAATGTATAGGTGAGATG 62803
Qy 274 erLeuLeuLeuGlnGlyLeuGluValGlyAspThrSerIleGlnGluTyrrIleIleG 294
Db 62802 CTTGTGTGGTGGAGGTCCTTGAATGTGAGTGAACAACATAGAGAAATATATATG 62743
Qy 294 lYrProAlaThrGlnAlaIleAspAspLeuValThrAspGlyAspLyGlnIleThrL 314
Db 62742 GTCCCGCGAGTAACTTACTGATGATCATGATATAGCTGTGATGAAGATCAATTAAGT 62683
Qy 314 euTyrrGlyProGlnGlnGlyGlnSerTyrrIleAlaArgProSerLyGlyProIleMetL 334
Db 62682 TATATGTCGACAGAGAGGCTGAGTGGTGTGCTAGCGCATCAAGGA----- 62633
Qy 334 euGlySerValLeuSerLeuAlaSerArgIleGlySer---MetValAsnGlnSerValP 353
Db 62632 --GGAAGACATAGATGATGTTGTGTCGCGCATGAGATCAATGACGAGGCAAGGCT 62575
Qy 353 roLeuMetAspProIleValThrLeuPheGlySerValHisGlnLysMetProGlnAlaG 373
Db 62574 CATATGATGATCTTGTCTACACTGTTTGGAGCGTTCACGAGAAATGCCGACT- 62516
Qy 373 lYgLySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspG 393
Db 62515 --GGAAGCATGAGAGTGGCTTGTCCCAATTTTGGAGTATGTTGATGTTGAGGGA 62458
Qy 393 lnhIleAlaLysArgGlnIntPrAspGlnGluAsnLeuHisArgAspArgGlyLysTy 413
Db 62457 ATCAACCAAGCATAGATGGAGTGAAGAAATCTTGTGGAAGAGTGGAGATTATC 62398
Qy 413 lAspAspGlyAlaGlyGlyAspTyrrGluAspAsnLeuHisSerProLeuSerArgArg 433
Db 62397 CATCCGAC---CATGAGATGATCTTGAAGATGATCTTCACTTCTGATGATCTCAAGT 62341
Qy 433 lnhAlaThrGlyAlaGlnGlyLysAspIleValHisIleGlyHisArgGlySerAlaLeuS 453
Db 62340 AAACGACAAAGCATGAG---AAAGACATGCTCACACATGCTCATGGAATCTTCTTCACT 62284

```

```

Qy 453 erMetArgArgGlnThrLeuLeuGlyGlyGlyAspGlyValSerSerThrAspIleG 473
Db 62283 TCACACATGGAAGTCAAGTGAAGAGCTCAAGGGGAAGAGCGGATGATGGGATTG 62224
Qy 473 lYgLyGlyTyrrPheGlnLeuAlaTyrPheSerGlyLysGlyGlyGluAsnGlyArgLySg 493
Db 62223 GAGGTGATGTCGAATGTCATGGAATGGAACGGAAGAGAAATGAAATCGGACAGAAAG 62164
Qy 493 lYgLyGlyPheLyArgValTyrrLeuHisGlnGlnGlyValProGlySerArgArgLyS 513
Db 62163 AAGGTGGTATTAACCGATATACTCTGATCAAGAGGTTTCCCGAGATCTCGAGTGGCT 62104
Qy 513 erIleValSerLeuProGlyGlyGlyAspValPheGlnGlySerGluPheValHisAla 533
Db 62103 CAATGTTTCACTTGTGCTGGTGTGATGGAACCGGTAG--GCAATTTTGTACAGGCT 62047
Qy 533 lAlaLeuValSerGlnSerAlaLeuPheSerLyGlyLeuAlaGluProArgMetSerA 553
Db 62046 CTGCTTGTGTTAGCCAAACAGCTCTTATTCAAAGACCTTCTCAAGAAACATACATTTG 61987
Qy 553 spAlaAlaMetValHisProSerGluValAlaAlaLyGlySerArgTyrrLyAspLeuP 573
Db 61986 GTCCGTGATGTATCATCCCA--ACACTAAAGGTCATTTGGCATGATCTTC 61930
Qy 573 heGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln---- 591
Db 61929 ATGATCTCGAGTCAAGCGGTGATTAAGTCGATGAGATTGACCTTCAATATCTTCAAGC 61870
Qy 591 ----- 591
Db 61869 TAGCTGATATGCAACATCAAAATTAATGAATCACAAGAAATATGCGTTATATGCTA 61810
Qy 592 -----GlnPheAlaGlyIleAsnGlyValLeuT 601
Db 61809 ATTCTATGTTTCTCGTCTCTTGGAAATGCAAGTCTCAGGATCAACGAGATCTTCT 61750
Qy 601 YrTyrrThrProGlnIleLeuGlnGlnAlaGlyValAlaValIleLeuSerTyrrPheGlyL 621
Db 61749 ACTACACACCGCAATCTTGAAGAGGCGGTGTGCGGATCTTATCTATCGAACAATGGGA 61690
Qy 621 euSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProC 641
Db 61689 TTAATCTTCTTCTGACATCTTACTTATTAAGTGAATGACAACTTGTATGTTACTG 61630
Qy 641 ysrIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg----- 654
Db 61629 CAATAGCTGTTCAATAGAGGCTCATGATCTTCTGTCGAAGGTGATTAACCTTAAT 61570
Qy 654 ----- 654
Db 61569 TTCCTTGTCTTACATATACATTAACGAACAGCTAAATGCTTATGAATGTCCTT 61510
Qy 655 -----ArgPheLeuLeuLeuGlyTyrrIleProIleLeuIleAlaSerLeu 669
Db 61509 TTATTTTATATTCAGAGACTTGTCTTCAACCGATCAATCTCGATAGCATCTCTA 61450
Qy 670 ValIleLeuValIleSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSer 689
Db 61449 TTGGTTTATGATATCTCAATCTTGTTCACATGAACAGCATGTGACGCGGTCTTATCA 61390
Qy 690 ThrValSerValIleValTyrrPheCysPheValMetGlyPheGlyProIleProAsn 709
Db 61389 ACCGTAAAGCTTGTGCTTACTTGTCTTGTGATAGGTTTGGTCTGCTCCAAAC 61330
Qy 710 lLeuGlyAlaGlnIlePheProThrArgValAlaArgGlyLeuGlyValIleAlaIleCysAla 729
Db 61329 ATCTCTGTTTGAAGATTTTTCACATCGAGTCCGCGGAATCTGATCGCATCTGGCA 61270
Qy 730 PheThrPheTyrrIleGlyAspIleIleValThrTyrrSerLeuProValMetLeuAsnAla 749
Db 61269 CTCACCTTGTGATCTGACATATATGTCATTACAGTCTCCCGTGTGCTCAAAATCC 61210

```

Oy	750	116g1ylem1acgiya1phaPseSellietyrAla1valCySeulleieserPhieValPhe	769
Db	61209	ATTGACATAGCGTGGTGTTGTGAATGACCAATCTTAATGTCATTCATCATGCGTCTTT	61150
Oy	770	ValPheLeulySValProgluThFlyeGlYmeProleugluValIierThruPiPePhe	789
Db	61149	GTCGTCATTAAAGTCGCCGAACATTAAAGCAGTCGACTTAGAGTCATCACAGGTCCTTT	61090
Oy	790	AlAvalGiYa1AlaySGInal1a1a1a1a	798
Db	61089	TCTGTTGAGCTAGACAAGCTGAGACT	61063
<hr/>			
RESULT 13			
LOCUS	AC136843	135406 bp	DNA linear PLN 06-JUN-2003
DEFINITION	Oryza sativa (japonica cultivar-group) chromosome 11 clone		
ACCESSION	OSUNBA0094M06,	complete sequence.	
VERSION	AC136843		
KEYWORDS	AC136843.4 GI:31442544		
SOURCE	HTG.		
ORGANISM	Oryza sativa (japonica cultivar-group)		
REFERENCE	Oryza sativa (japonica cultivar-group)		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartidae; Oryzaceae; Oryza.		
REFERENCE	1 (bases 1 to 135406)		
AUTHORS	Bharгава,A., Dalai,V., Sharma,T.R., Pal,A.K., Singh,A., Yadav,M., Gaikwad,K., Srivastava,S., Dixit,A., Ghazi,I.A., Batra,K., Mohapatra,T. and Singh,N.K.		
TITLE	Complete sequence for Oryza sativa chromosome 11 clone OSUNBA0094M06		
JOURNAL	Unpublished		
AUTHORS	2 (bases 1 to 135406)		
REFERENCE	Swain,S.C., Gaikwad,K., Pal,S., Yadav,M., Sureshababu,K., Ghazi,I.A., Singh,A., Srivastava,S., Dixit,A., Bharгава,A., Pal,A.K., Dalai,V., Batra,K., Sharma,T.R., Mohapatra,T. and Singh,N.K.		
TITLE	Direct Submission		
JOURNAL	Submitted (08-NOV-2002) IIRGS, NRC on Plant Biotechnology, Indian Agricultural Research Institute, LBS Centre, New Delhi, Delhi 110012, India		
REFERENCE	3 (bases 1 to 135406)		
AUTHORS	Swain,S.C., Gaikwad,K., Pal,S., Yadav,M., Sureshababu,K., Ghazi,I.A., Singh,A., Srivastava,S., Dixit,A., Bharгава,A., Pal,A.K., Dalai,V., Batra,K., Sharma,T.R., Mohapatra,T. and Singh,N.K.		
TITLE	Direct Submission		
JOURNAL	Submitted (28-FEB-2003) IIRGS, NRC on Plant Biotechnology, Indian Agricultural Research Institute, LBS Centre, New Delhi, Delhi 110012, India		
REFERENCE	4 (bases 1 to 135406)		
AUTHORS	Bharгава,A., Dalai,V., Sharma,T.R., Pal,A.K., Singh,A., Yadav,M., Gaikwad,K., Srivastava,S., Dixit,A., Ghazi,I.A., Batra,K., Mohapatra,T. and Singh,N.K.		
TITLE	Direct Submission		
JOURNAL	Submitted (06-JUN-2003) IIRGS, NRC on Plant Biotechnology, Indian Agricultural Research Institute, LBS Centre, New Delhi, Delhi 110012, India		
COMMENT	On Jun 6, 2003 this sequence version replaced gi:28604240.		
FEATURES			
SOURCE	Location/Qualifiers		
	1..135406		
	/organism="Oryza sativa (japonica cultivar-group)"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:39947"		
	/chromosome="11"		
	/clone="OSUNBA0094M06"		
BASE COUNT	39875 a	28472 c	27393 g 39665 t 1 others
ORIGIN			
<hr/>			
Alignment Scores:			
Pred. No.:	3.15e-146	Length:	135406
Score:	2368.00	Matches:	492
Percent Similarity:	73.22%	Conservative:	104

Query	DB	Similarity	Matches	Mismatches	Indels	Gaps
US-10-051-909-32 (1-800) x AC136843 (1-135406)		60.44%	58.11%	156	62	20
19	ArgAAGGluAArgLeu-----LeuProSerValValLeuAlaLeuProGlyPro	34				
99031	CGACCCACACGGGTGATGAAACCCAAACACCCCGCTTGGTTGACACAAATCATCCA	99090				
35	LeuProProAla-----SerCysSerSerGlnGluProValThrSerAsp	49				
99031	ATCCCTTCGCGCTTTTCATGTGGTTCTCTGAGAAACCATCCAGATC-----	99141				
50	AspIleLeuGluAspLysMetSerGlyAlaValLeuValAlaIleValAlaSerIleGly	69				
99142	GATCGATCGCCACGTAACGATGAGAGCGCGCGTGGTGGTGGCGCGCGCCCATCGGC	99201				
70	AsnLeuLeuGlnGlyTTPAspAsnAlaThrIleAlaAlaValLeuIleValIleValS	89				
99202	AACTACCTGACAGGATGGAGAACACCAACATCGCGCGCGCGTGTGTAACAAAGCGG	99261				
90	GluPheGlnLeuGluAsnGluProThrValGluGlyLeuIleValSerMetSerLeuIle	109				
99262	GAGTTTCGCCCTCGAAGACCAAGCCCGCGTGGAGGCGCTGTGTGCTGCATGTCTCTATC	99321				
110	GlyAlaThrIleValThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgPro	129				
99322	GCGCGCACCATCATCAACACCTTTCGCGCGCGCGCTCGCATCGTGGCGCGCGCC	99381				
130	MetLeuIleLeuSerSerIleLeuTyrrPhePheSerGlyLeuIleMetLeuTTPSerPro	149				
99382	ATGCTCATCGCT	99441				
150	AsnValTyrrValLeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaVal	169				
99442	AAGCTCAAGCT	99501				
170	ThrLeuValProLeuTyrrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsn	189				
99502	ACCTCTGCTCCCGCT	99561				
190	ThrLeuProGlnPheSerGlySerGlyIleMetPheLeuSerTyrrCysMetValPheGly	209				
99562	ACGCTGCGCGCAAGTTCACCGCGCTCGCGCGCAATGTCATCTCTCTCTCTCTCTCTCT	99621				
210	MetSerLeuSerProSerProAspThrArgIleMetLeuGlyValLeuAlaIleProSer	229				
99622	ATGACGCTCTGCGCT	99681				
230	LeuPhePhePheGlyLeuThrIlePheTyrrLeuProGluSerProArgThrLeuValSer	249				
99682	CTGCTCTACCTGTTGTCACCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	99741				
250	LysGlyIleArgMetAlaGluAlaLysValLeuGlnLysLeuArgGlyLysAspAspVal	269				
99742	AAGGCGCCCATGAAAGAGGCGCGGCTGTGCTCGAGATCTGCGCGCGCGAGAGCTC	99801				
270	SerGlyIleLeuSerLeuLeuGlnGlyIleGlnGlyValGlyIleValIleValIleVal	289				
99802	TCCGCGAGATGCGCT	99861				
290	GluTyrrIleIleGlyProAlaThrGlnAlaIleAspAspLeuValThrAspGlyAspLys	309				
99862	GACTACGCTCTGCGCGCT	99903				
310	GluGln-----IleThrLeuTyrrGlyProGlnGlnGlyIleValIleValIleVal	325				
99904	GAGCAGCGAGAGGACACGCTACGCTGTACGCGCGAGACAGAGGCTTCTGTGTGTGCG	99963				
326	ArgProSerLysGly-----ProIleMetLeuGlySerValLeuSerLeu--AlaSer	342				
99964	CAGCGGTGTCGCGCGCGCGCGCGAGCAGCTGTGGAGAGCTCTCTGTGGGCTGTGACG	100023				


```

QY 343 ARGHISGLYSERMETVALASNGINSEVALPROLEUMETASPROILEVALTHIRLEUPHE 362
DB 100024 CGCCATGGCAGCATGTACAGACAG-----ATGAAGACACCCCGTGTGGCGCTTTC 100074
QY 363 GLYSERVALHIGLUSNMETPROGLNALAGLY-----GLYSERMETARG--- 377
DB 100075 GGGAGCGTCCACGAGCGCGTGGCGAGTCCGGCGCGGCCACCGCGCAGCATGAGGGGG 100134
QY 378 SERTHIRLEUPHEPROASNPHEGLYSERMETPHESERVALTHIRASPNHIALALYSAN 397
DB 100135 AACACGCTGTCCCACTCGGAGACATGCTTAAGCTCAACGATAGGCCCGGC---GGC 100191
QY 398 GIUGINTPRASPGLUGLUSAN-----LEUHSARGASAPSPGL 410
DB 100192 AACACGTGGAGAGAGAAACGTGCAGCTGGCGACGACGACTCGACAGAGAGAGAG 100251
QY 411 GLUTYRALASER---ASPGLYALAGLYASPTRYGLUASPASNLEUHSERPROLEU 429
DB 100252 GAGTACCTCTCCGACGACGCGCAAGAGACGACGACGCGCGCGCTGCAGGACACCGCTG 100311
QY 430 LEUSERARGLINALAETHIRGLYALAGLYLYSASPILAEVALHIBIGLYHISARGGLY 449
DB 100312 CTGTGCGGCGACAGACCCGACGTCGAGCAAGAACAGACCCCGCTCCGCGCAG----- 100365
QY 450 SERALALEUSERMETARGARGGLINTHIRLEUAGLYGLUGLYLY-----ASPGLYVAL 467
DB 100366 -----GTGGCATGACAGCGGACGACGACGACGCGCGCGCGCGCGCTCGACAGCGGG 100419
QY 468 SERSERTHIRASPILEGLYGLYLYTRYPGLNEUALATPRLYSERTSPERGILULYSGLUGLY 487
DB 100420 AGCAGCATGCGCATGCGCGCGGCGGTGCGAGCTGCGCTGGAGATGGAGAGACGAGGCG 100479
QY 488 GIUANGLYARGLYGLUGLYLYPHELYSARGVALTYRLEUHIAGLN-----GLUGLY 505
DB 100480 CCGGACGCGGTCAACGCGCGCGCGGTGAGACCCGATGTACTCGACGAGAGTCCAGAGCC 100539
QY 506 VALPROGLYSERARGARGGLYSEIRILEVALSERLEUPROGLYGLYVAPVALPHEGLU 525
DB 100540 GCGCGCGCGGTGATCTCGGCGCC-----GCCGCGACGCG---CAG 100578
QY 526 GLYSERGLUPHEVALHISALALALALEUVALSERGINSEVALAUEPHESEIRLYSGLY 545
DB 100579 TCCATGCGGCGTCAACGCGCGCGGTGAGCGGTGATGCGGTGATGATGATGATGATGATG 100638
QY 546 ---LEUALAGLUPROHGMETSERASPAALAMETVALHISPROSEIRLU----- 561
DB 100639 GTCTGATCGGCGAGAGCCCGACGCGCGGTTCGCGACACCGCGCGGAGCGGTGCGG 100698
QY 562 ---VALALALALYSEIRARGTRYLSAPLEUPHEGLUPROGLYVALARGARGALA 580
DB 100699 GCGGCGCGGTCAACGCGCGCGCGCGGTGCGAGCTGCTCGAGCGCGCGCTCGCACGCGC 100758
QY 581 LEULEUVALGLYVALGLYILEGLINILEUGINGINPHEALAGLYLEASNGLYVALLEU 600
DB 100759 CTCTTCTGCGGCGTCAACATCCATCCAGCAATTCCTCGGATCAACGCGCGTCTC 100818
QY 601 TYRTRYTHIRPROGLINILEUGLUGLINALAGLYVALALAVAILLEUSERLYSPHEGLY 620
DB 100819 TACTACACCCCGCAGATCTCTCAACGAGCGCGGTGATGAGGTCTCTCTCGCACGCTCGGC 100878
QY 621 LEUSERSERALASERLASEIRILEUULESERSETLEUTHIRLEUUMETLEUPRO 640
DB 100879 CTCTCGGCGACTTCACTCTCATCTCAACGAGCGCTCAACAGCTCTCTCAAGCTCCCG 100938
QY 641 CYSILEGLYPHEALAMETLEUUMETASPLEUSERGLYARGARGPHELEUULEUGLY 660
DB 100939 TCCATGCGGCGTCAACGCGCGTCAACGAGCTGCGCGCGCGCGCGCTCTCTCTCTCG 100998
QY 661 THIRLEPROILEUULEALASERLEUVALILEUVALVALSERASNLEUULEASPLEU 680
DB 100999 AGCTGCGCGTCTCTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 101058

```

```

QY 681 GLYTHIRLEUALAHIALALEULEUSERTHIRVALSERVALILEVALTYRPHCYSCYSPE 700
DB 101059 GCGGCGCGCGGCGAGCGCGCGCTGTCAACGCGGAGCGGTCACTGTCTGTCTGTCT 101118
QY 701 VALMETGLYPHEGLYPROILEPROASNILEUCYSALAGLUILEPHEPROTHIRARGVAL 720
DB 101119 GTCATGCGGCGTCCGCGCGCATCCCAACATCTCTGCGCGAGATCTTCCACAGCGGGTG 101178
QY 721 ARGGLYLEUCYSILEALALECYALAEHPHERPHEPRILEGLYASPILEVALHIR 740
DB 101179 AAGGAGCTCTCATTTGCGATCTGCTGCTGAGCTTGTGGCTGGCGGACATGCGCTCAAG 101238
QY 741 TYRSEIRLEUPROVALMETLEUASNALALEGLYLEUALAGLYVALPHESEIRILETYRALA 760
DB 101239 TACAGCTCCCGCTCATCTCAAGCTCGGTGGGCGCTCGCGCGGTCTCTCTTCTACGCC 101298
QY 761 VALVALCYSEULESEIRPHEVALPHEVALPHELEULYSVALPROGLUTHRLYSGLYMET 780
DB 101299 GCCGTGTCTGCTGCGCGCGCTGTGTGCTGCGCTCAAGGTGCGCGAGACCAAGGCGCTC 101358
QY 781 PROLEUGLVALILETHIRGLUPHEPHEALAVAGLYVALYS 794
DB 101359 CCGCTCGAGGTATCATCATGATCTTCAACGTGCGCGCCAGG 101400
RESULT 14
AX412656 2205 bp DNA linear PAT 14-JUN-2002
LOCUS AX412656
DEFINITION Sequence 420 from Patent WO0222675.
ACCESSION AX412656
VERSION AX412656.1 GI:21445114
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (chale crese)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 Glazebrook, J., Wang, X., Dangl, J. L., Bulgem, T. and Zhu, T.
Plant genes, the expression of which are altered by pathogen
infection
Patent: WO 0222675-A 420 21-MAR-2002;
SYNGENTA PARTICIPATIONS AG (CH) ; UNIVERSITY OF NORTH CAROLINA AT
CHAPEL HILL (US) ; GLAZEBROOK, Jan (US) ; Wang, Xun (US) ; Dangl,
Jeffrey L. (US) ; Bulgem, Thomas (US)
FEATURES
source
location/qualifiers
1..2205
/organism="Arabidopsis thaliana"
/mol type="genomic DNA"
/db xref="taxon:3702"
BASE COUNT 521 a 524 c 521 g 639 t
ORIGIN
Alignment Scores:
Pred. No.: 1..3e-142 Length: 2205
Score: 2275.00 Matches: 462
Percent Similarity: 75.89% Conservative: 111
Best Local Similarity: 61.19% Mismatches: 146
Query Match: 55.83% Indels: 36
Gaps: 14
US-10-051-909-32 (1-800) x AX412656 (1-2205)
QY 56 METSERGLYVALLEVALALALEVALASERILEGLYASNLEUENGINGLYTRP 75
DB 1 ATGAAGGAGGAGCACTCTGTGCTCTCGCCCGCAATCGCAATTCTTCAAGAGATGG 60
QY 76 ASPANALATHIRILEALALALEVALLEUITYRILEULYSGLUPHEGLINLEUASN 95
DB 61 GACATATGCCACATTTGCTGAGCTATGTTATATCAACAAAGACTTAATCTA----- 114
QY 96 GIUPROTHR---VALGLUGLYLEUILEVALSERMETSERILEUILEGLYVALTHIRLEVAL 114
DB 115 ---CCAACCTCTGTTCAAGGTCTTGTGCTGTCTATGTATGATGAGCGTGAACGGTATTC 171

```

115 ThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleuSer 134
 172 AGGACTGTGTCAGAGCCGATATCTGATGGCTCGGAGAGCGCCCATGCTCTTTATCA 231
 135 SerIleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeu 154
 232 TCAGTATGATATTTGCTGCGGTTTGATATGTTATGTTGCTCCCAATGCTATGTTCTG 291
 155 LeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeu 174
 292 TGCCTTGCAGGCTCTTAATAGGTTTGCTGCGGAGCTCGCGTTACACTTCTCCCTGT 351
 175 TyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPhe 194
 352 TAGATTCTGAAACCGCTCTCCGAGATCAGAGGACACTTAATACCTCTCCCTCAAGTTT 411
 195 SerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerPro 214
 412 CTGGCTCTGGGATGTTTGTGATCTGTATGTTTCACTATGCTCCTGAGTGAC 471
 215 SerProAspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGly 234
 472 TCCTCTACCTGAGAGCCATGCTCGAGTCTCTCTGATCCCTCTCTTTATTTGTTT 531
 235 LeuThrIlePheTyrLeuProGluSerProArgTyrPheValSerIleGlyArgMetAla 254
 532 CTCAGCGGTTTATTTGCTCCGAGTCTCTCTGCTGCTGTTAGTAAAGAAAGATGAC 591
 255 GluAlaValLeuValLeuGluIleLeuArgGlyValAspAspValSerGlyLeuLeuSer 274
 592 GAGGCTAAGCCGAGTTCTTCAACAGTTATGTCAGAGAAAGATGTTACCATGAGATGCT 651
 275 LeuLeuLeuGluGlyLeuGluValAlaGlyAspThrSerIleGluGluTyrIleLeuGly 294
 652 TTACAGTGTGAAGGCTGATATAGAGAGAAAGAAACATGAGATCTCTTACTACT 711
 295 ProAlaThrGluAlaAlaAspAspLeuValThrArgGlyAspLeuGlu---GlnIleThr 313
 712 TTGAGAGATCATGAGATGATGATACCTGTAACCTGTTAGAGATGAGACAAATGGCG 771
 314 LeuTyrGlyProGluGluGlyGlnSerTyrIleAlaArgProSerIleGlyProIleMet 333
 772 CTTATATGAGACCCAGAGATCAATCACTGTCTGACCTGTCCAGAACAA----- 825
 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
 826 ---AATAGCTCACTGGCTACGCTCTCGCCACGAGACTTACAAACCAAGCATGTATC 882
 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
 883 CTTAAGATCCGCTCGTCAATCTTTTGGCACTCCACAGAAATATCCAGAACACAGG 942
 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
 943 GGAACACTCGGAGGAGATTTTCCCTCATTTCCGAGAGATGTTCACTACTACCTCCGAT 1002
 394 -----HisAlaValAsnGluIleTyrAsp-----GluGluAsnLeuHisArgSer 408
 1003 GCGCTCTCAGGTAACCGGCTCATTTGGAAGAAAGACATAGAGAGCATTAACAACAAAGC 1062
 409 AspGluGluTyrAlaSerAspGlyAlaGlyAspTyrGluAsp-----AsnLeu 425
 1063 AATATGACTATGCACTGATGATGTCGGGTGATGATGATGATGATGATGATGATGATGAT 1122
 426 HisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyValAspIleValHisHis 445
 1123 CGTAGCCCTTAATATGTCGCGCACAGACACAGCATGAC---AAGATATATGATCCCAT 1179
 446 GlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyValAsp 465
 1180 CCTCAACAGTGAAGACATTTAAGCATGAGACACAGTACGATTATGCAA---GGCAAC 1236

466 GlyValSerSerThrAspIleGlyGlyTyrPheGluLeuAlaTyrPheTyrSerGlyLys 485
 1237 GCGGAAGTATACATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1296
 486 GluGlyGluAsnGlyValGlyGlyGlyGlyPheLysArgValTyrLeuHisGlnGly 505
 1297 GAA-----TACAAAGAGTATTAATTAAGAAAGATGGA 1329
 506 ValProGlySerArgArgGlySerIleValSerLeuProGlyGlyValAspValPheGlu 525
 1330 CTTGAA---TCTCCCGGCTGATCATCTTAATTCCTGAGAGGTCGCGAT-----GCT 1380
 526 GlySerGluPheValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSerGly 545
 1381 GGAGGAGCTATCATTAACGCTTCGCTTGTATGAGAGATGTTGTTCTTGTGCTTAATCA 1440
 546 LeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLys 565
 1441 -----GTTCAATGATCCGCGCATGCTCCCGGAGAAATGCTGCTCT 1485
 566 GlySerArgTyrAspLeuPheGluProGlyValAlaArgArgAlaLeuValGlyVal 585
 1486 GGACCACTGTGCTGCTCTTCTTGAACCTGTGTAAAGCTGTGCTGTGCTGTGCTGTGCT 1545
 586 GlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGln 605
 1546 GGCATTCAAAATCGCAGAGATTTTCAAGTATCATAGAGATCTCTACTACACCTCTAG 1605
 606 IleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSer 625
 1606 ATTTCTGCAACCGGCTGCGTGTGATATCTTCTTCCGAGCCTCGAGCATTAAGTTCCATCTCT 1665
 626 AlaSerIleLeuIleSerSerLeuThrThrLeuMetLeuProCysIleGlyPheAla 645
 1666 GCGTATCTTCTCATGAGGCTTATTAACAATTAATTAATTAATTAATTAATTAATTAATTA 1725
 646 MetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGluGlyThrIleProIleLeu 665
 1726 ATGAGACTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1785
 666 IleAlaSerLeuValIleLeuValAlaSerAsnLeuIleAspLeuGlyThrLeuAlaHis 685
 1786 ATTTCTCATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1845
 686 AlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGly 705
 1846 GCAGCACTCTCCAGAGT 1905
 706 ProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIle 725
 1906 CCATTTCAAAACATCTCTGTTTGAATCTTCCCAACAGAGCCGAGTCTGTGATC 1965
 726 AlaIleCysAlaPheThrPheTyrIleGlyAspIleIleValThrTyrSerLeuProVal 745
 1966 GCCATATGTGTATGCTTGTGATTTGAGACACTTATTTGACAGTACTCACTTCCGCTT 2025
 746 MetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValAlaCysLeuIle 765
 2026 CTCTCACTGATGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2085
 766 SerPheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIle 785
 2086 TCATGATCTTCTGTTTACATGAAGTCCCGAGACTAAAGCATGCTTTGGAGATATC 2145
 786 ThrGluPheAlaValAlaGlyAlaLysGlnAlaAlaLysAla 800
 2146 ACAAGTACTTGTGCTTGGAGCT---CAAGCTCAAGCTTCTGCT 2187

RESULT 15
 AX507559
 LOCUS AX507559 2205 bp DNA linear PAT 27-SEP-2002
 DEFINITION Sequence 2254 from Patent WO0216655.
 ACCESSION AX507559

[illegible]

Db 592 GAGGCTAAGCCAGATTCTTCAACAGTATATGTGGCAGAGAAAGATGTTAACCGATGACATGGCT 651

Qy 275 LeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGly 294

Db 652 TTACTAGTTGAAGAGCATGATATATAGGAGAGAAAAACAATGGAAGATCTCTTAGTAAC 711

Qy 295 ProIatThrGluAlaIleAspLeuValThrAspGlyAspGlu---GlnIleThr 313

Db 712 TTGGAGATCATGAAGGTGATGATACACTGTGAACCGTTGATGAGATGACAAATTCGG 771

Qy 314 LeuTyrGlyProGluGluGlyGlySerTrpIleAlaArgProSerIleGlyProIleMet 333

Db 772 CTTTATGGAAACCCAGAGAAATCATCTGATCTTGTAAGCTGTGCCAGACA----- 825

Qy 334 LeuGlySerValLeuSerLeuAlaSerArgIleGlySerMetValAsnGlnSerValPro 353

Db 826 ---AATACCTACCTGGGCTACGCTCTGCCACGGAGCTTATGCAACCAAGCATGATC 882

Qy 354 LeuMetAspProIleValThrLeuPheGlySerValIleGluAsnMetProGluAlaGly 373

Db 883 CTTAAAGATCCGCTCTGTCATCTTTTGAGCTTCCACGAGAAATGTCAGAAAGAGCGC 942

Qy 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393

Db 943 GAAACACTCGGAGTGGGATTTTCCCTCATTTTCGGAAGCATGTTCAAGTACTGCGCAT 1001

Qy 394 -----HisAlaIleAsnGluGlnTrpAsp-----GluGluAsnLeuHisArgAsp 408

Db 1003 GCGCCTCACGGTAAACCGGCTCATTTGGGAAAAGACATAGAGAGCATTAACAACAAGAC 1061

Qy 409 AspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAsp-----AsnLeu 425

Db 1063 AATGATGACTATGCGACATGATGATGGTGGGTGATGATGATGATGATGATGATGATG 1122

Qy 426 HisSerProLeuLeuSerArgGluAlaThrGlyAlaGluGlyLysAspIleValHisHis 445

Db 1123 CGTACCCCTTAAATGTCGCGCCAGCCACAAGCATGAC--AAGAGTATGATCCACAT 1177

Qy 446 GlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyIleAsp 465

Db 1180 CCTACAAATGGAAGACATTAAAGCATGAGACACACAGTACAGCTTATGAA---GGCAAC 1233

Qy 466 GlyValSerSerThrAspIleGlyGlyTyrGlnLeuAlaTrpLysTrpSerGluLys 485

Db 1237 GCGCAAAATGATGATGGAATGGTGGTGGTGCATATGCGATATGATATGACAAAACGAT 1291

Qy 486 GluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGly 505

Db 1297 GAA-----TACAAAGGTATTAATCTTAAAGAAATGGA 1332

Qy 506 ValProGlySerArgArgGlySerIleValSerLeuProGlyGlyLysValPheGlu 525

Db 1330 GCTGAA---TCTCGCGGTGGCTCGATCATCTTATTCGCGAGGCCGAT-----GGT 1388

Qy 526 GlySerGluPheValHisAlaAlaIleValSerGlnSerAlaLeuPheSerIleGly 545

Db 1381 GGAGCGAGCTACATTCACGCTTCGCGCTGTGAACAGATCTGTTCTTGCTCTTAATCA 1441

Qy 546 LeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLys 565

Db 1441 -----GTTCATGATCGCGCAAGGTTTCCCGCGAGAAATTCGCTCT 1488

Qy 566 GlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyVal 585

Db 1486 GGACCACTCTGGTCTGCTCTTCTTAACCTGGATGAACGCTGCTGGTGTGGTGTGC 1541

Qy 586 GlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGln 605

Db 1546 GGCATTCAAAATATCGACAGATTTCAAGTATCAATGAGTTCTCTAATACTACCTCTCAG 1601

Qy 606 IleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSer 625

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 18:39:20 ; Search time 41 Seconds
(without alignments)
1876.462 Million cell updates/sec

Title: US-10-051-909-32

Perfect score: 4075

Sequence: 1 IRSGSWLAQVPPFTDLDLR.....PLEVITFEFVAGAKQAAKA 800

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2584	63.4	729	2 T06127	probable sugar tra
2	2278	55.9	734	2 H86340	Sugar transporter
3	2245	55.1	729	2 T45780	sugar transporter
4	2183	55.6	734	2 T51139	sugar transport pr
5	587.5	14.4	457	2 E70070	metabolite transpo
6	536	13.2	580	2 D86426	hypothetical prote
7	533	13.1	580	2 D70073	metabolite transpo
8	524.5	12.9	461	2 D84772	probable sugar tra
9	520.5	12.8	582	2 F71431	hypothetical prote
10	518.5	12.7	469	2 H97064	probable sugar-pro
11	505	12.4	473	2 G69789	sugar transporter
12	493.5	12.1	508	2 G84564	probable sugar tra
13	493	12.1	493	2 A85433	sugar transporter
14	489	12.0	521	2 G84864	probable membrane
15	488	12.0	549	2 T14606	probable sugar tra
16	486.5	11.9	482	2 B69803	metabolite transpo
17	474	11.6	547	2 C84593	probable sugar tra
18	473.5	11.6	464	2 F69587	L-arabinose transp
19	465	11.4	419	2 B69888	metabolite transpo
20	462	11.3	612	2 B40538	myo-inositol trans
21	460.5	11.3	511	2 A84537	probable sugar tra
22	459.5	11.3	511	2 H84536	probable sugar tra
23	458	11.2	560	2 T51485	sugar transporter-
24	451	11.1	464	2 F65079	galactose-proton s
25	448	11.0	464	2 C91106	galactose-proton s
26	448	11.0	464	2 F85951	galactose-proton s
27	445	10.9	464	2 AC0877	galactose-proton s
28	435.5	10.7	471	2 AB0868	L-arabinose isomer
29	434.5	10.7	472	2 S47089	arabinose-proton s

30	434	10.7	491	2 A26430	xylose transport p
31	434	10.7	491	2 F91255	xylose-proton symp
32	434	10.7	491	2 B86096	xylose-proton symp
33	432.5	10.6	490	2 T14545	probable sugar tra
34	432	10.6	557	2 T38125	myo-inositol trans
35	431	10.6	547	2 A48442	membrane transport
36	429.5	10.5	472	2 B26430	L-arabinose isomer
37	429.5	10.5	472	2 B91091	L-arabinose isomer
38	429.5	10.5	472	2 B59936	L-arabinose isomer
39	421.5	10.3	584	2 B69555	myo-inositol trans
40	418.5	10.3	468	2 S10014	glucose transport
41	417.5	10.2	555	2 S69671	hypothetical prote
42	415	10.2	606	2 T27072	hypothetical prote
43	412	10.1	487	2 B96782	hypothetical prote
44	412	10.1	502	2 B70845	probable sugar tra
45	407.5	10.0	613	2 T27077	hypothetical prote

ALIGNMENTS

RESULT 1

T06127
probable sugar transport protein P23E12.140 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 22-Oct-1999
C:Accession: T06127
R:Byvan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Hohnsels, submitted to the Protein Sequence Database, April 1999
A:Reference number: Z15485
A:Accession: T06127
A:Molecule type: DNA
A:Residues: 1-729 <BEV>
A:Cross-references: EMBL:AL022604; GSPDB:GN00062; ATSP:P23E12.140
A:Experimental source: cultivar Columbia; BAC clone P23E12
C:Genetics:
A:Gene: ATSP:P23E12.140
A:Map position: 4
A:Introns: 27/1; 216/1; 433/1; 519/3; 582/2
C:Keywords: sugar transporter; transport protein

Query Match	Score	Length	DB 2;	Length	DB 2;
Best Local Similarity	68.8%	Pred. No. 9.7e-174;			
Matches: 512;	Conservative: 89;	Mismatches: 123;	Indels: 20;	Gaps: 8;	
QY	56	MSGAVLVAIVASIGNLQGMNATIAAIVLYIKKEPOLQNEPTVEGLIVMSLIGATIVT	115		
DB	1	MSGAVLVAIAAVGNLLQGMNATIAAGAVLYIKKEFNLSNPVEGLIVAMSLIGATIT	60		
QY	116	TFSGPLSDISGRRLPMLTISLTFPSGLIMTMSPPVYVLLARFVGGIGLAATLVPLX	175		
DB	61	TCSGGVADWLRRLPMLTISLTFPSGLIMTMSPPVYVLLARFVGGIGLAATLVPLX	120		
QY	176	ISEIAPSRIGLNTLTPQSSGGMFSLCYCWFQMSLSPSPDWRIMLGVLAIPSLFFPGL	235		
DB	121	ISEIAPSRIGLNTLTPQSSGGMFSLCYCWFQMSLSPSPDWRIMLGVLAIPSLFFPGL	180		
QY	236	TIFVLPESPRVLVSKGRNAEAKVYLQKRGKDVSGEISLLLEGLVEVGGTISIEYITGP	295		
DB	181	TIFVLPESPRVLVSKGRNAEAKVYLQKRGKDVSGEISLLLEGLVEVGGTISIEYITGP	240		
QY	296	ATEAADLVITGQDEKQITLVGPEGGQSWIARPSGPIMLGVLASLASHGS-MNQSVP	354		
DB	241	ADEVTDHDIADVDDQIKLYAEERGLSVVAPVKG---GSTVSLVSHSGTMSRRQSL	296		
QY	355	MDPIVTLFGSVHKNMPQAGSGMRSTLFPNFGSMFSTVDQAHKNBQWDEBNLHRDDEEYAS	414		
DB	297	IDPLVTLFGSVHKNMPD-GSMRSALFPNFGSMFSTVDQAHKNBQWDEBNLHRDDEEYPS	355		
QY	415	DGAGGVDYDNLHSPILSRQATGAEKDIIVHGHRSALSMRQITLLGGGQGVSTDTIGG	474		
DB	356	D-HGDDSDDLHSPILSRQATGAEKDIIVHGHRSALSMRQITLLGGGQGVSTDTIGG	413		

QY 475 GQWLAKMSKEGKRGKGFYVYLHOEGVPGSRGSIIVSLPGGADVFEGBEFVHAA 534
 DB 414 GQVAMKMTERRDEGGQKE-----EGFPGSRGSIIVSLPGGDDGGE-ADVQASA 462
 QY 535 LVSGQALPSKGLAEPMSDAAMVHSEVAAKSRWMDLFEPCVRRLALVGVGIQILQOFA 594
 DB 463 LVSGQALPSKGLAEPMSDAAMVHSEVAAKSRWMDLFEPCVRRLALVGVGIQILQOFS 521
 QY 595 GINGVLYTPQILBOAGVAVILSKFGLSASASIIISITLTLMMPCIGFALMLDLSGR 654
 DB 522 GINGVLYTPQILBOAGVAVILSKFGLSASASIIISITLTLMMPCIGFALMLDLSGR 581
 QY 655 RELIGTIPILIASVILVAVNSLIDGLTAHALLSVSYIVYFCGFWGFGPIPNILCAE 714
 DB 582 RLTLTLTIPILIASVILVAVNSLIDGLTAHALLSVSYIVYFCGFWGFGPIPNILCAE 641
 QY 715 IFTVRGLCIAICAFTEWIGDIIVTYSIPVMLNAGLAVFSIYAVVCLISFVFLVKV 774
 DB 642 IFTVRGLCIAICAFTEWIGDIIVTYSIPVMLNAGLAVFSIYAVVCLISFVFLVKV 701
 QY 775 PETKMPLEVITEPFAVGAQKAA 798
 DB 702 PETKMPLEVITEPFAVGAQKAA 725

RESULT 2

H86340
 Sugar transporter [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: H86340
 R:Theologian, A.; Becker, J.R.; Palm, C.D.; Federici, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 Ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiz, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Talion,
 Ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.
 A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: H86340
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-734 <STO-
 A:Cross-references: GB:AE005172; NID:g4836905; PIDD:AA030608.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1

Query Match 55.9%; Score 2278; DB 2; Length 734;
 Best Local Similarity 61.2%; Pred. No. 3.3e-152;
 Matches 462; Conservative 112; Mismatches 145; Indels 36; Gaps 14;
 QY 56 MSGAVLVAIVASIGNLQGMNATIAAVALYIKKEFOLONEPT-VEGLIVNSLIGATV 114
 DB 1 MKGATLVAAATIGNFLQGMNATIAAGAVYINKOLN---FTVQGLVAVMSLIGATV 57
 QY 115 TFSGLPSDSIGRRPMLIISILYFPGSLIMLSPNVYLLARVVDGFGIGLATVLYPL 174
 DB 58 TFSGLPSDSIGRRPMLIISILYFPGSLIMLSPNVYLLARVVDGFGIGLATVLYPL 117
 QY 175 YISELAPSRIRGLNTLTPQFSSGGMFLSYCMVFCGMSLSPDNRIMGVLAIPSLFFPG 234
 DB 118 YISELAPSRIRGLNTLTPQFSSGGMFLSYCMVFCGMSLSPDNRIMGVLAIPSLFFPG 177
 QY 235 LTFILPSPRMLVSKGRMAEAKVYLQKLRGKDVSGELSLLEGLVGVGDSIEEYIIG 294
 DB 178 LTFILPSPRMLVSKGRMAEAKVYLQKLRGKDVSGELSLLEGLVGVGDSIEEYIIG 237
 QY 295 PATEADLVTDGKE-QITLVGPEGQSWIARPSKPIMLGSLVSLSRHSGMNVQSV 353
 DB 238 LEDHGDGDTLFTVDEGQWRLTYGTHENGSLARPVBEQ---NSSIGLSRHSGLANQSMI 294

QY 354 LMDPIVTLFGSVHNMPOAGSMRSTLFPNFGSPSVTDQ--HAKNEQWD---EENLRD 408
 DB 295 LKDPVNLFGSLHMKPAGGNTSGIFPHFGSSSTTADAPGKPAHMEKDISHTYKD 354
 QY 409 DEEYASDAGDGYED---NLHSPILSRQATGAEKDIYVHGRSALSMRQTLLEGCD 465
 DB 355 NDVATDDGACDDDDSDMDLNLPLMSRQTSM-DKMIPIHPTSGTSLMRHSTLMQ-GN 412
 QY 466 GVSSTDIQGGQWLAKMSKEGKRGKGFYVYLHOEGVPGSRGSIIVSLPGGADVFE 525
 DB 413 GESSWGIGGGHWMGYRYENDE-----YKRYLLEKDGAE-SRSGIISIIPGPD--G 460
 QY 526 GSEFVHAALVSQALPSKGLAEPMSDAAMVHSEVAAKSRWMDLFEPCVRRLALVGV 585
 DB 461 GSYIHSAVSALYSKGLAEPMSDAAMVHSEVAAKSRWMDLFEPCVRRLALVGV 515
 QY 586 GIQILQOPAGINGVLYTPQILBOAGVAVILSKFGLSASASIIISITLTLMMPCIGFA 645
 DB 516 GIQILQOPAGINGVLYTPQILBOAGVAVILSKFGLSASASIIISITLTLMMPCIGFA 575
 QY 646 MLMDLSGRFLLGTIPILIASVILVAVNSLIDGLTAHALLSVSYIVYFCGFWGFG 705
 DB 576 MLMDLSGRFLLGTIPILIASVILVAVNSLIDGLTAHALLSVSYIVYFCGFWGFG 635
 QY 706 PIPNICAIEFTVRGLCIAICAFTEWIGDIIVTYSIPVMLNAGLAVFSIYAVVCL 765
 DB 636 PIPNICAIEFTVRGLCIAICAFTEWIGDIIVTYSIPVMLNAGLAVFSIYAVVCL 695
 QY 766 SFVFLVKVPEPKMPLEVITEPFAVGAQKAA 800
 DB 696 SFVFLVKVPEPKMPLEVITEPFAVGAQKAA 729

RESULT 3

T45780
 Sugar transporter-like protein - Arabidopsis thaliana
 N:Alternate names: protein F26013.130
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
 R:Deisen, M.; Berger, C.; Cooke, R.; Grellet, F.; Landie, M.; Mewes, H.W.; Lemcke, K.;
 submitted to the Protein Sequence Database, December 1999
 A:Reference number: Z23013
 A:Accession: T45780
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-729 <DEU-
 A:Cross-references: EMBL:AL133452
 A:Experimental source: cultivar Columbia; BAC clone F26013
 C:Genetics:
 A:Map position: 3
 A:Introns: 27/1; 216/1; 523/3; 578/2
 A>Note: F26013.130

Query Match 55.1%; Score 2245; DB 2; Length 729;
 Best Local Similarity 60.4%; Pred. No. 6.8e-150;
 Matches 457; Conservative 110; Mismatches 144; Indels 46; Gaps 14;
 QY 56 MSGAVLVAIVASIGNLQGMNATIAAVALYIKKEFOLONEPTVEGLIVNSLIGATV 115
 DB 1 MRSVIVLVAALAINMQLGMNATIAAGAVYIKKEFHEKPKIEGLIVMSLIGATIT 60
 QY 116 TFSGLPSDSIGRRPMLIISILYFPGSLIMLSPNVYLLARVVDGFGIGLATVLYPL 175
 DB 61 TFSGLPSDSIGRRPMLIISILYFPGSLIMLSPNVYLLARVVDGFGIGLATVLYPL 120
 QY 176 ISELAPSRIRGLNTLTPQFSSGGMFLSYCMVFCGMSLSPDNRIMGVLAIPSLFFPG 235
 DB 121 ISELAPSRIRGLNTLTPQFSSGGMFLSYCMVFCGMSLSPDNRIMGVLAIPSLFFPG 180
 QY 236 TIFPLPSPRMLVSKGRMAEAKVYLQKLRGKDVSGELSLLEGLVGVGDSIEEYIIG 295
 DB 236 TIFPLPSPRMLVSKGRMAEAKVYLQKLRGKDVSGELSLLEGLVGVGDSIEEYIIG 295

Db 181 AAFLLPESPRMLVSKGRNDEARQVLRGRGVEDSGELALVEGLGVGKDTISIEBYVIGP 240

Qy 286 ATEAADLVLDGDKKQITLYGPEEGSWTARPSKCPIMLGSVLSLSRHHGSVNVNSVPLM 355

Db 241 DNEBEGNEELPRKQIKLYGPEEGSWAKKPVKQ---SLSLASRSGSLPFGSILM 296

Qy 356 DEIVTLFGSVHNMNMP--QAQGSMTSLFPNFGSMFSTVDQAKNQMDEBNLHRDDEEYA 413

Db 297 DPLVTLFSGIHENLPSSENNASRSMFLPNNMSIIGMGR--QESQWDEP---RNED-- 349

Qy 414 SDGAGDYEDNLHSLSRQATGAAGKDI VHHGHSALSMRRQTL-LGEGDGVSTDI 472

Db 350 ---SSDDENLNSLPQT--EPDD--YHQRVTGTHRRQSSLFMANVGETATATSI 401

Qy 473 GGGWOLAMKSEKEENGRK--EGFKRYVYHOE-----GVPSRRGSIVSL--GGGDV 523

Db 402 GGGWOLAMKYNKVDGKRVNGVQORMYIHEETANNNTNTPFRRRGSLSFHEGGQH 461

Qy 524 FEGSEFVHAALVSQSALFSKGLAEPMSDAMVHPSEVAAGSRWKDLFEPCVRALLV 583

Db 462 DVNVGVQAALVQSASMMRGKGTAML-----PKEV-KDGPGRRELKEGVRALMV 514

Qy 584 GVGIOILQOPAGINGVLYTTPQILEQAGVAVILSKFGLSSASASIISSLTLLMPCIG 643

Db 515 GVGIOILQOPAGINGVLYTTPQILEETGVSSLTLMGISASBSASIISSLTLLMPCIT- 573

Qy 644 FAMLMDLSGRRLGTPPIILASVILVNSLIDGLTAHALSTVSYVYVFCGFVWG 703

Db 574 -----LVMSKSLMSTPIILISLVTLVLSLVNLSGSLNALSSTASVTVYVLSGFWG 626

Qy 704 FGPINILCAEIPTRVRCICIAICAFPMIGDIIVTYSLPVNLNAGLAGVSIYAVVC 763

Db 627 FGAINILCEIFPSVSGLCITTCITCALPTMCDIIVTTLPMKLSIGLAGVGIYAVVC 686

Qy 764 LISFVFLKVPETKMPLEVITEFPFVAGAKQAANA 800

Db 687 AVAWFVFLKVPETKMPLEVISEFVSAGAKQDAAA 723

RESULT 4

T51139
sugar transport protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 28-Jul-2000 #sequence_revision 28-Jul-2000 #ext_change 28-Jul-2000
C:Accession: T51139
R:Jaden, J.; Neuhans, E.
submitted to the EMBL Data Library, August 1995
A:Description: A new sugar transport protein from Arabidopsis thaliana.
A:Reference number: Z25311
A:Accession: T51139
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-734 <TUA>
A:Cross-references: EMBL:Z50752; PIDD:CAA90628.1

Query Match 53.6%; Score 2183; DB 2; Length 734;
Best Local Similarity 59.5%; Pred. No. 1.6e-145;
Matches 449; Conservative 110; Mismatches 160; Indels 36; Gaps 14;

Qy 56 MSGAVLVAIVASIGMLQGMNATIAAVALYIKKEFQONEPT-VEGLIVSSILGATTV 114

Db 1 MGATLVLAATIGNFGLOGMNATIAAGANVYINKDLN---PTVSGLVVAMSLIGATYI 57

Qy 115 TTFSGPLSDIGRRPMLLISLILYFSGILMLSPNVYLLARFVDGFGIGLAVTVPL 174

Db 58 TTCGSPISDMLGRRPMLLISVYVFCGILMLSPNVYLLCFARLLNGFGAGLAVTVPV 117

Qy 175 YISEIAPSEIRGLINTLPQSSGGMELSYCWVFGMSLSPSDWRIMGLVAILPSLFFG 234

Db 118 YISEIAPSEIRGLINTLPQSSGGMELSYCWVFGMSLSPSDWRIMGLVAILPSLFFG 177

Qy 235 LTIIFYLSPRWLVSKGRMAEAKVLOKRGKDDVSGELSLLEGLGVGDTISIEBYIG 294

Db 178 LTFEYLPESPRMLVSKGRNDEARQVLRGRGVEDVTDEMLVLEGLDIGEKTEMDLLVT 237

Qy 295 ATEAADLVLDGDKKQITLYGPEEGSWTARPSKCPIMLGSVLSLSRHHGSVNVNSVPLM 353

Db 238 LEDHEGDDTLTVEDEGQIRLYGTHENOSTYARVPEP---NSSLGRSRHSGSLANSMT 294

Qy 354 LMDDIVTLFGSVHNMNMPQAQGSMTSLFPNFGSMFSTVDQ--HAKNQMDE--BNLHRD 408

Db 295 LKDPVLVNLFGSLHMKMPAGNTSGIFPHHGSFSTTADAPKQPAWKEIDISHVKD 354

Qy 409 DEEYASDAGGDYED---NLHSLSRQATGAAGKDI VHHGHSALSMRRQTL-LGEGD 465

Db 355 NDYATVDGADDDDDSDNDLRSPLMSRQTSMD-KDMI PHPTSGSTLSMRHSTLMQ-GN 412

Qy 466 GVSSTDIQGGQOLAMKSEKEENGRKRGKRYVYHOEYVPSRRGSIVSLPFGGDVFE 525

Db 413 GESSMGIGCGWHMGVRYENDE-----YKRYLLKEDGE--RRRGSIIISPGPD--G 460

Qy 526 GSEFVHAALVSQSALFSKGLAEPMSDAMVHPSEVAAGSRWKDLFEPCVRALLV 585

Db 461 GGSYTHASALVSRLVGRPKS-----VHGSANVPBKILASGLMSALLEPVRKALVGV 515

Qy 586 GIQILQOPAGINGVLYTTPQILEQAGVAVILSKFGLSSASASIISSLTLLMPCIGFA 645

Db 516 GIQILQOPAGINGVLYTTPQILERAGVDILSLSLGSLISASFLISGLITLLMPCIVVA 575

Qy 646 MLLMDLSGRRLGTPPIILASVILVNSLIDGLTAHALSTVSYVYVFCGFVWG 705

Db 576 MRLMDVSRRLSLMTPIILVLSLVTLVLSLHISKRVNALSIGCVLVLCFFVWYG 635

Qy 706 PIPNILCAEIPTRVRCICIAICAFPMIGDIIVTYSLPVNLNAGLAGVSIYAVVC 765

Db 636 PFOSSVYKSSQOARDGICIAICAFPMIGDIIVTYSLPVNLNAGLAGVSIYAVVC 695

Qy 766 SFVFLKVPETKMPLEVITEFPFVAGAKQAANA 800

Db 696 SWIFVFLKVPETKMPLEVITDYPAFGA-QAQASA 729

RESULT 5

E70070
metabolite transport protein homolog ywcG - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #ext_change 20-Jun-2000
C:Accession: E70070
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter, C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Galler, Jech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M.F. Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinis, A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y. M.; Ogawa, K.; Ogilwara, A.; Oudega, B.; Park, S.H.; Parvo, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, A:Authors: Schleich, S.; Schroeder, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror, Akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K. A:Authors: Yoshikawa, H.F.; Zumsche, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: E70070
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-457 <KUN>
A:Cross-references: GB:Z99122; GB:AL009126; NID:92636029; PIDD:CAB15600.1; PID:92636109
A:Experimental source: strain 168
C:Genetics:

Query Match 14.4%; Score 587.5; DB 2; Length 457;
Best Local Similarity 22.8%; Pred. No. 1.4e-33;
C:Superfamily: glucose transport protein

Matches 168; Conservative 102; Mismatches 162; Indels 305; Gaps 12;

QY 55 KMAGVALVAIVASIGNLQGMNATIAAVALYIKKEFOQNEPTVEGLIVSLIGATIV 114
 DB 2 KKQSIMLYFFGALGALYDVTGIVSGALYFMKELG--NAFEGLVSSLLGALIG 59
 QY 115 TTFSGPLSDSIGRRPMLILSYFFPSGLIMLSPVYVLLARFVDFGIGLATVLYPL 174
 DB 60 SGAAKLTDRFGKKAIMAAALLFCIGGLVALAPNTGVMLFRIILGLAVGTSITVPL 119
 QY 175 YISLAPSEIRGLNTLPQSGSGMFLSYCMVFGMSLSPSDMRIMGLVLAIPSLFPFG 234
 DB 120 YLSLAPKHKRGALSLNQLMTVIGLLSY--IVVTFADAEAMWMGLAIVPDLIL- 176
 QY 235 LTFPLPSPRMLVSGKMAEAKVLOKRGKDVSGELSLLEGLVGGDTSIEYIIG 294
 DB 177 IGLTFPSPRMLFTNGEESKAKKILDKRGTKD----- 211
 QY 295 PATEAADLVTDDGEQITLYGPEBQSWIARPSKPIMLGSLSLASHRGSVMNQVPL 354
 DB 212 ----- 211
 QY 355 MDPIVTLFGSVHENDPQAGSMRSTLPNFGSMFSVTDQAHAKNEQMBENLRHDEBYAS 414
 DB 212 ----- 214
 QY 415 DGAGDYEDNLHSPILSRQATGAEKDIYHHGRGALSMTROTLLGEGGDVSTDIG 474
 DB 215 ----- 218
 QY 475 GWOLAMKSEKENGKRGKRVYLLHOGVPSGRSIVSLPGGDVFEGBEFVHAA 534
 DB 219 -----KEAEKQDEGGLKE----- 231
 QY 535 LVSQALSKGLAEPRMSDAMVHPSEVAAGSRKMDLPEPGVRALLVGVGIQLQOPA 594
 DB 232 -----LFDPMVRPALIAGLAEFLQOFI 254
 QY 595 GINGVLYYTPOLLEQAGAVILSKFGLSASATL-----ISLTLTLMPCIGPMLMD 650
 DB 235 GNTTIIYAFKFTTVG-----FG--NSASILCTVIGITVNNLMTL---VAIKIID 300
 QY 651 LSGRRPILLTGTPILIASLIVLVNSLIDLTALHALLSTVSIVYFCFCWFGPPIPI 710
 DB 301 KIGRRPILLFAGNAGVILVALLVNLFPNNTPAASMTVICLGPIVFAVNSWGPVAV 360
 QY 711 LCAELFPFRVGLCIACTFWIGDIYVTSLPVNLNAGLAVPSIYAVCLISFVAV 770
 DB 361 MLPELFPPLHVRGIGVSTLMLHVGTLVSLTYPIIMEAIGISYFLIYAAIGMAFLV 420
 QY 771 FLKVBETKGMPLVITE 787
 DB 421 RPKVTEYTKGRSLLEIEQ 437

RESULT 6
 D86426
 hypothetical protein fl2p21.2 - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cross)
 C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002
 C/Accession: D86426
 R/Theologian: A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewart, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Lin, X.; Lin, S.X.; Liu, Z.A.; Luos, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Xu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.
 A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A/Reference number: A86141; MUID:21016719; PMID:11130712
 A/Accession: D86426

A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-580 <STO>
 A/Cross-references: GB:A8005172; NID:g11120774; PIDN:AAG30955.1; GSTDB:GN00141
 C/Genetics:
 A/Map position:
 C/Superfamily: 1
 Query Match 13.2%; Score 536; DB 2; Length 580;
 Best Local Similarity 21.6%; Pred. No. 8.4e-30;
 Matches 175; Conservative 93; Mismatches 169; Indels 372; Gaps 11;

QY 62 VAIVASIGNLQGMNATIAAVALYIKKEFOQNEPTVEGLIVSMIGATIVTTESGP 120
 DB 31 LAPASIGGLIFGVDVIGSALYIYDDPFSVDRNTMLQMIYMAVAGLVAGALIGW 90
 QY 121 LSDISGRPMLILSYFFPSGLIMLSPVYVLLARFVDFGIGLATVLYPLISEIA 180
 DB 91 ANDKIGRSALIMADPFLGAIIMAAAFNPSSLVGVGFVGLGVMASTAPLYISEAS 150
 QY 181 PSEIRGLNTLPQSGSGMFLSYCMVFGMSLSPSDMRIMGLVLAIPSLFPFGITLYPL 240
 DB 151 PAKTIGALVSTNGFLTGQGLSYLINAFI-DVTGTRMMLGIAGLPALIQFVL-MFTL 208
 QY 241 PESPRMLVSKGRMAEAKVLOKRGKDVSGELSLLEGLVGGDTSIEYIIGPATEAA 300
 DB 209 PESPRMLYRKREBEAKAILRIYSADVEGEIRALQDSVT----- 250
 QY 301 DDLVTDGKQITLYGPEBQSWIARPSKPIMLGSLSLASHRGSVMNQVPLMDPIVT 360
 DB 251 -EILEBGSSEKINMI----- 264
 QY 361 LFGSVHENDPQAGSMRSTLPNFGSMFSVTDQAHAKNEQMBENLRHDEEYASDAGGD 420
 DB 265 ----- 264
 QY 421 YEDNLHSPILSRQATGAEKDIYHHGRGALSMTROTLLGEGGDVSTDIGGMOLAW 480
 DB 265 -----KICKAT----- 271
 QY 481 KWSEKENGKRGKRVYLLHOGVPSGRSIVSLPGGDVFEGBEFVHAAALVSOSA 540
 DB 272 ----- 271
 QY 541 LFSKGLAEPRMSDAMVHPSEVAAGSRKMDLPEPGVRALLVGVGIQLQOPAGINGVL 600
 DB 272 -----VRGLIAGVGLVFOQFVGINTVM 295
 QY 601 YTPPOLLEQAGAVILSKFGLSASATLISLTLTLMPCIGPAMLMDSGRFPILIG 660
 DB 296 YSPFTIVQLAGFA-----SRTALLSLVTAAGNAPGSIISYFIDRIGKKLLII 346
 QY 661 TTPILIASLVTL-----VNSLIDLTG----- 682
 DB 347 SLFGVILISGLITGVFEAATHAPALISLETORFNNISCPDYKSAAMTNAMDCTULKAS 406
 QY 683 -----LAH-----ALSTVSIVLY 696
 DB 407 SPSCGYCGSPIGKEHPGACWISDSYKDLCHNENRMLYTRGCPSPNFGMFALLGGLYIIF 466
 QY 697 FCCFVWGPGPIPNILCAEIPFRVRGICIACTFWIGDIYVTSLPVNLNAGLAVGF 756
 DB 467 FS-----PGGIVPMLVNSEIYPLRRGICGIAAIAANNISNLIYVQSFLEALIGTSWTF 523
 QY 757 SIYAVVCLISFVFLKVBETKGMPLVITE 785
 DB 524 LIFGVISVIALFLVMVVCVPEETKGMPEBI 552

RESULT 7
 D70073
 metabolite transport protein homolog yxoc - Bacillus subtilis
 C/Species: Bacillus subtilis

[illegible]

Db 313 GSVITITSLAALSGV - LITLGLSASTAMKTVVFLGVIYVFYATGAPVWVWMLPELFS 370

QY 719 RVRLGIALCAFTFWIGDITVITSLPVMNALGIAGVFSIYAVVCHISFVFLKVPETK 778

Db 371 KARAANGFTTIVLSAANLIVSLVFLPMLSAMGIAMFWFVSICLSFFFAVYWPETK 430

QY 779 GMPLEVT 785

Db 431 GKSLLEI 437

RESULT 8

Db4772

Probable sugar transporter [imported] - Arabidopsis thaliana

CSpecies: Arabidopsis thaliana (mouse-ear cress)

Cdate: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002

CAccession: DB4772

R.Li, X.; Kaul, S.; Rounsley, S.D.; Shee, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffitt, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.; Nuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999

ATitle: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

ATReference: A84420; MUID:20083487; PMID:10617197

AAccession: DB4772

AStatus: Preliminary

A.Molecule type: DNA

A.Residues: 1-580 <STO>

A.Cross-references: GB:AE002093; NID:G4263781; PIDN:AMD15441.1; GSPDB:GN00139

C.Genetics:

A.Gene: Atg35740

A.Map position: 2

C.Superfamily: glucose transport protein

Query Match 12.9%; Score 524.5; DB 2; Length 580;

Best local Similarity 20.9%; Pred. No. 5,4e-29;

Matches 171; Conservative 102; Mismatches 175; Indels 369; Gaps 13;

Db 62 VAIVASIGNLLOGWDMNTIAAVALYIKER-QLONEPTEGLIVSMGLIGATVITTPSGP 120

QY 29 LALASAGIGLIFGNTGVINAGLLIYKEEGERVDNKLQELIIVSMVACAIYGAALIGW 88

Db 121 LSDSISRRPMLISSLIYFFSGLIMTSPVNVYLLARFVDGEGIGLATVTLVPLIYSEIA 180

QY 89 YNDFGRMSVLTADVFLGLGALVMVIAHPWYIIIGRLIVGCGVAMSTSPLYISEMS 148

Db 181 PSEIRGLNLTLPQSSSGCMFLSCWFCGMSLSPSPDMRMGLVLAIPSLFFGLTFYVL 240

QY 149 PARIRGALVSTNGLLITGGPFLSLINLAFVHPG-TWRMMLGVSALPAIIQCL-MLTL 206

Db 241 PSEPRWLVSKRMALAKVLOKLRGKDVSGEISLLEGELEGVGDTSIEEYIIGPATEAA 300

QY 207 PSEPRWLVSKRMALAKVLOKLRGKDVSGEISLLEGELEGVGDTSIEEYIIGPATEAA 300

Db 301 DLVLTGDKQITLYGPEGQSWIARPSKGPIMLGSVLSLASHRHGSNVNOSVPLMDPIVT 360

QY 252 DB----- 253

Db 361 LFGSVHNMPPAGGSMRSTLFPNFGSMFVTTDQAKRQWDEENLRHDEEYASDDAGCD 420

QY 254 ----- 253

Db 421 YEDNLSPLSRQATGAEGKDIYVHGRGSALSMRQTLTGGEGDGVSTDIGGQOLAW 480

QY 254 -----DIICH----- 258

Db 481 KWSKEGNGRKEGFKRVYILHDEGVPSRGSTIVSLPGGADVFESEFVHAALVYSQA 540

QY 259 TFSRK-----LRGA----- 267

Db 541 LFSGLAEPKMSDAAMHPSEVAKSGRMWDLEPGVRAALVGVGQILQGFAGINGVL 600

QY 268 -----LSNPVRHGLAAGITIVQVAGQFVGIVTWM 236

Qy	601	TTTPOILEAGAAVAVILSRFGSSASATLISLTLMPCIG--FAMLLMDLSGRRL	658
Lp	237	YSPFLLDPAGYA-----SNTKATLALITS-----GLNAVGSVSNMFDRYGRKLM	345
Qy	659	LGTTPIILASIVILV-----	673
Db	346	IISMGGIITCVILAAVFNESANHPKIDKDSRNFAGKATCPAPAFPTASRSPSNWC	405
Qy	674	-----VSNLIDLGLLHA-----LSTVSIV	695
Db	406	MKCLQYDCGFCSSNGAQEVAPGACTVQASADMALCHSKRTPFKQGCSPGYLAVFLGL	465
Qy	696	YFCFCVMGFGLPNILCAEIEPTRVRGCIACAFTEFWIGDIIYVSLPVLNAIGLAV	755
Db	466	YIIVAPGMGTVPWIVNSEIYVPLRYRGLAGGIAAVSNMNSMLVYSEFLTLITLNAVSSGT	525
Qy	756	FSIYAVVCLISFVFLKVPETKMPLEVIITFEFVVG	792
Db	526	FLIFAGSSAVGLFTFMLLVPETKGLQFEVEVEGLTEGG	562

RESULT 9

hypothetical protein - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
A/Variety: Columbia
C/Date: 03-Aug-1998 #sequence_revision 03-Aug-1993 #text_change 20-Jun-2000
C/Accession: F71431
R/Beran, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
P.; Weller, H.; Wedler, E.; Wanduth, R.; Weltzienegger, T.; Pohl, T.M.; Terry, N.; Gie
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A/Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erhoff, A.; Moore, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; An
C.; Chalvarria, N.
A/Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
A/Reference number: A71400; MUID:96121113; PMID:9461215
A/Accession: F71431
A/Status: Preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-582 <BBV>
A/Cross-references: GB:297341; NID:g2244991; PID:g2245004
C/Genetics:
A/Map position: 4COP9-4G3845
A/Superfamily: glucose transport protein

Query Match	12.8%	Score 520.5	DB 2	Length 582
-------------	-------	-------------	------	------------

Best Local Similarity 20.3%; Pred. No. 1e-28;
Matches 166; Conservative 106; Mismatches 180; Indels 365; Gaps 12

OY 62 VAIIVASIGNLLQGMWNATIAAIVLTIKKEF-OLONIPVTEGLIVMSLIGATIVTFSPG 120
::: :
DB 30 IAIISAGTGGILLFGYPTGVTSICALPIIKEDNDREVDKRTWLOSTIVSNMAVAARIVGAACGW A9

121 LSDSIGRRPMLISSILYFPGSLIMLMSPNVYVLLARFVDGEGIGLAVTLVPELYISETA 180

Db 90 INDKPGRMSILADVLFLIGATVAPAPAPWIIIGRIFVFGVGMSMTSPLEYISEAS 149

```

181 PSEIRGLINTLTPFGSGGEMFLSYCMVFGMSLSPSIDWRIMCVLAIPSLFFPGLTIFYL 240
      ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
190 PARPFCYUVCWMCITTECCOPEEYIYRATNNEEDMAYTQACVAGCATHOTNHYHIC 207
      ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :

```

241 PESPRUUVSKGMAEAKKYLQKLRGKDVSGELSLILEGLVGGDTSEEBYIIGPATEA 300

```
Db      208 PESPRLYRKDRIAESRAILRIYPADVEAEAMEALKLSVE-----AEKA 252
```

```

0y 301 DdLVtDGDKEQITLLYGPBEGQSWIARPSKGPIMLGVLSLASRHGSMVNQSVPLMDPIVT 360
    |:::
    |:::
    |:::

```

DB	253	DEA1	-----	256
OY	361	LFSGVHEHMPQAGGSMRSTLF	PNFGSGMFSVTDHAI:NEQWDEENLHRDDEEYASD	GAGED 420

—

Db 257 -----IGDSFSA----- 263

421 YEDNLHSPLLSROATGAEGKDI VHHGHRGSAISMRRQTLLGEGGDVSTDI GGGWQLAW 480

	-	:				
D _b	-----	XU	KGA	F	G	----
	264					270

QY	481	KMSEKEENGKRGKGFKRIVYLHQEGVPGSRGSI	IVSLPGGQDVEEGSEFVHAALVSGSA	540
Ph	271	-----	-----	270

541 LPSKGLAEPKMDAAMVHPSEVAKGRWKDLFEPGYRALLVGVGIQILQDFAGINGVL 6000

Db 271 -----NPVVRGLAGITVQVAQGFVSGINTVM 297

```
Dh      ||| :||| :  
601 YTPQIEAGAVILSKGLSSASAILISLTLMLPCTIGRAMLLMDLSGRFILLG 660  
Dh      ||| :||| :  
298 VCSPTIOEGV-----SNKEMAMISLTETCI NAIYGT -----VSMEFDNDVCPPIMTI 348
```

238	115FSLVVFSLA-----VSNLIDL	378
239	SNLPLPLDLSLLDGLPLDGL	
240	-----VSNLIDL	680
241	661 TIPILASLVIL-----VSNLIDL	
242	-----VSNLIDL	
243	-----VSNLIDL	
244	-----VSNLIDL	
245	-----VSNLIDL	
246	-----VSNLIDL	
247	-----VSNLIDL	
248	-----VSNLIDL	
249	-----VSNLIDL	
250	-----VSNLIDL	
251	-----VSNLIDL	
252	-----VSNLIDL	
253	-----VSNLIDL	
254	-----VSNLIDL	
255	-----VSNLIDL	
256	-----VSNLIDL	
257	-----VSNLIDL	
258	-----VSNLIDL	
259	-----VSNLIDL	
260	-----VSNLIDL	
261	-----VSNLIDL	
262	-----VSNLIDL	
263	-----VSNLIDL	
264	-----VSNLIDL	
265	-----VSNLIDL	
266	-----VSNLIDL	
267	-----VSNLIDL	
268	-----VSNLIDL	
269	-----VSNLIDL	
270	-----VSNLIDL	
271	-----VSNLIDL	
272	-----VSNLIDL	
273	-----VSNLIDL	
274	-----VSNLIDL	
275	-----VSNLIDL	
276	-----VSNLIDL	
277	-----VSNLIDL	
278	-----VSNLIDL	
279	-----VSNLIDL	
280	-----VSNLIDL	
281	-----VSNLIDL	
282	-----VSNLIDL	
283	-----VSNLIDL	
284	-----VSNLIDL	
285	-----VSNLIDL	
286	-----VSNLIDL	
287	-----VSNLIDL	
288	-----VSNLIDL	
289	-----VSNLIDL	
290	-----VSNLIDL	
291	-----VSNLIDL	
292	-----VSNLIDL	
293	-----VSNLIDL	
294	-----VSNLIDL	
295	-----VSNLIDL	
296	-----VSNLIDL	
297	-----VSNLIDL	
298	-----VSNLIDL	
299	-----VSNLIDL	
300	-----VSNLIDL	

Db 349 SMFGITACIIITLVFSQAIHAPRIDAFESRTFAPNATCSAYAPLAAENAPPSRWNCMK 408

```
Qy      681 -----GLTAAH-----LISTSVIVPF 697
```

DB 409 CURSECGFCASGQPIAPGACVLDSDMKATCSSRGRFEDGCGPSKFGDLAVFMDLYI 468

QY 698 CCFWMGFGPIPNILCAEIEFPTRVRGICIAICAFTFWIGDIIVTYSLPVMLNATIGLGVFS 757

Db 469 VYAPMGCTPVIWNSSEYPLRYRGIGGIAAVSNWVSNLIVSEFSLSTHALGSSGTF 528

QY 758 IYAVVCLISFVFEVLKVPETKGMPLFVITEPFAVGAK 794

RESULT 10

H97064 probable sugar-proton symporter (imported) - *Clostridium acetobutylicum*

Cispecies: Clostridium acetobutylicum
CDate: 14-Sep-2001 #sequence 14-Sep-2001 #text 30-Sep-2001
CRevision: 107064

R.Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, C. *Accession*: 55/054

Article: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium *Clostridium acetobutylicum* 183, 4823-4838, 2001

A,Reference number: A96900; MUID:21359325
A,Accession: H97064

A, Status: preliminary
A, Molecule type: DNA

A;Residues: 1-469 <KOR>
A;Cross-references: GB:AE001437; PID:AAK79307.1; PID:g15024271; GSPDB:GN00168

a:Gene: PAC139
 b:Experimental source: CLOSTRIDIUM ACETODUPLICUM ATCC824
 c:Genetics:

C:Superfamily: glucose transport protein

Query Match	12.7%;	Score 518.5;	DB 2;	Length 469;
Best Local Similarity	21.8%;	Pred. No. 1.1e+28;		

Matches 164; Conservative 94; Mismatches 175; Indels 319; Gaps 12;

60 VLVAVIASIGNULOGMDNATIAAATLYIKKEFOLONEPTWEGGLVSMSLIGATVITTSFG 119

15 VLISCAGIGLELYDPAVISAIGLEKLYNL--SBAMQFVISSIMVGVLGVSFG 72

0x F5D576KRCFNNLSDILIF7SGJNLMWSPNV1VUDDKAF VDDG9GTGVNF ILVFAL15BBI 1/5
120 :
Db FEGDAICBPBKTM AAAI PAISAVISSSTSSAPMMIPAAITVGGLGITGMASAI SVTYITREC 112
73

180 APSEIRGLNLTLPQFSSGGNLFSLYCMFQM-----SISPSDWRIMLGVAISLFFF 233

Db 133 APPSIRGRUSSLVQLFTILGISITPEFVNCIVNMGGSETRVSGMGRMLACTVAIVFL 192

234 GILTI¹FLPESPRWLVS²KGRMAEAKVYLQ³KLRGKDDVSGEISL⁴LEGLFVGGD⁵DSIEEYII 293

Db 193 -ITTFVESPRFLVKSNIKAAVLTAKINGABIAKQELDISKSLATENDSSD----- 246
 Qy 294 GPATBAADDVTDGKEQITTLXGPEGQSWIARPKSGIMLGVSILASRSGSMVQSV 353
 Db 247 -----GQ----- 248
 Qy 354 LMDPIVTLFGSVHNNPQAGSMRSTLPNFGSMFSVTDQAKNQWDEENLHRDDEYA 413
 Db 249 ----- 248
 Qy 414 SDGAGDYEDNLHSPILSRQATGAEKDIYHHGRGALSMMRQTLIGEGDGVSSDTIG 473
 Db 249 ----- 248
 Qy 474 GGMQLAMWSEKENGKRGKRYVYHQEGVPSRSGIVSLPGGDVFESEFYANA 533
 Db 249 ----- 248
 Qy 534 ALVSGSALFSKGLAEPMSDAMVHPSEVAAGSRMKDLFEPGVRRALLVGVIQILOQF 593
 Db 249 -----LDPGLRALLIGIFLAFNQA 270
 Qy 594 AGINGVLYTTPQILEQAGVAVILSKGLSASASITL--ISSITLLMLPCIGFAMLLMDL 651
 Db 271 IGMSITTYGPEIPQMIGF-----KNSSFLATSVIGVVEVSTIL-----AMFLIDX 318
 Qy 652 SGRRL-LGTTPIIASVILVNSLIDGLTAHALISTVSIVT-VYCCCFPMGF-----G 705
 Db 319 LGRKRLMEIGSAMAVFMILLI-----GTSFYIKLSNGFVILIFICFVSPCISMG 369
 Qy 706 PIPNLCAEIPFTRRGCIACAFPMIGDIIIVYSLPVMUNALIGLGVSEIYAVCLII 765
 Db 370 PIPMIDELIPNHLRAATGATITFLMGAMNALIQOFTPLMNGIGAYTFMFCGINVI 429
 Qy 766 SEVFEVLKVPETKGMPLVITEFFPAVGAQOA 797
 Db 430 CFLVTTKVPETKNSLSEIEKFWIPKSKONA 461

RESULT 11
 669789
 sugar transporter homolog ydJk - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
 C:Accession: G69789
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berber, C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Erttington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Huillo, M.F.; Koeter, P.; Koningsstein, G.; Krog, S.; Kumano, M.; Kurita, K.; Lapidus, S.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Mesuda, S.; Mausel, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetle, Rieger, M.; Rivola, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, A.; Authors: Scheich, S.; Schroeter, R.; Scoffone, F.; Sekiuchi, J.; Sekowska, A.; Seroy, A.; Authors: Schleich, S.; Tanaka, A.; Tanaka, T.; Tapestre, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wpat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K. A:Authors: Yoshikawa, H.F.; Zumestein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:98044033; PMID:9384377
 A:Accession: G69789
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Residues: 1-473 <KUN>
 A:Molecule type: DNA
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: ydJk
 C:Superfamily: glucose transport protein
 Query Match 12.4%; Score 505; DB 2; Length 473;

Best Local Similarity 22.3%; Pred. No. 9.6e-28;
 Matches 168; Conservative 106; Mismatches 176; Indels 302; Gaps 16;
 Qy 54 DKMGAVLVAIVASIGULQGMNATIAAVLYIKKEFQLONEPTVGLVSMIGATI 113
 Db 6 NQMSFLRTIILVSTFGGLFGYDGVNGALPYGEBDQNLNFTBGLVSSILFGAL 65
 Qy 114 VTTSGLPLSDIGRRPMLISSILYFPGSLMLSPVNYVLLARFVDFGIGLAVTLVP 173
 Db 66 GAVFGGRSDPNGRKNILFLAVIFFISTICTAPNTWITISRFVLGIVGASVTP 125
 Qy 174 LYISEIAPSEIRGLNLTLPQPSGSGMFLSYCM--VEGMSLP--SPDMRIMLVLAIPSL 230
 Db 126 AYLAEMSPVESRGMTQNELMIVSGQLAFVFAAII GTTMGNSHWRFVLVLAIPAL 185
 Qy 231 F-FPGITITFYLPSEPRMLVSKGRMAEKYLOKRGQDVDSGELSLBGLFVGDDISIE 289
 Db 186 FLFFGM--IRMPESPRMLVSKGRKEDALRYLKRR----- 218
 Qy 290 EYIIGPATBAADDVTDGKEQITLYGPEGQSWIARPKQIMLGVSILASRSGSMVN 349
 Db 219 ----- 218
 Qy 350 QSVPLMDPIVTLFGSVHNNPQAGSMRSTLPNFGSMFSVTDQAKNQWDEENLHRD 409
 Db 219 -----D 219
 Qy 410 EYVSDGAGDYEDNLHSPILSRQATGAEKDIYHHGRGALSMMRQTLIGEGDGVSS 469
 Db 220 EKRAA-----AELOEI----- 230
 Qy 470 TDIGGMQLAMWSEKENGKRGKRYVYHQEGVPSRSGIVSLPGGDVFESEF 529
 Db 231 -----EPAFKK----- 236
 Qy 530 VHAALVQSALFSKGLAEPMSDAMVHPSEVAAGSRMKDLFEPGVRRALLVGVIQI 589
 Db 237 -----EDGLEKAT-----FDLSVPMWRIRVIFIGLAI 265
 Qy 590 LQOPAGINGVLYTTPQILEQAGVAVILSKGLSASASITL--ISSITLLMLPCIGFAMLLM 649
 Db 266 VQOITGVNSIMYGEIILRNSG---FQTEALLININANGVIVLATFV-----GIWLL 315
 Qy 650 DLGRR-FLIGTPIIASVILVNSLIDGLTAHALISTVSIVYVCEFWNG-FGPI 707
 Db 316 GRVGRPMLMTGLTGTALLLIGIFSLVLE-GSPALPYV-VLSLTVFLAFQOGASPV 373
 Qy 708 PNILCAEIPFTRRGCIACAFPMIGDIIIVYSLPVMUNALIGLGVSEIYAVCLISF 767
 Db 374 TWIMSEIFPRLKNGLGCVVFCMMVWNPVAVSTFPILLALIGLSTTFIFVGLGICSV 433
 Qy 768 VFVFLKVPETKGMPLVITEFFPAV---GAKO 795
 Db 434 LFVKKFLPETKGLSLEQLEENFRAYDHSKAKK 465

RESULT 12
 664564
 probable sugar transporter [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C:Accession: G64564
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Beito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Kodiac, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayan, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A64420; MUID:20083487; PMID:10617137
 A:Accession: G64564
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-508 <STO>
 A:Cross-references: GB:A6002093; NID:94218010; PIDN:AAD12218.1; GSPDB:GN00139

C:Genetics:
A:Gene: At2g18480
A:Map position: 2
C:Superfamily: glucose transport protein

Query Match 12.1%; Score 493.5; DB 2; Length 508;
Best Local Similarity 22.7%; Pred. No. 6,8e-27;
Matches 170; Conservative 99; Mismatches 174; Indels 305; Gaps 19;

```

QY 63 AIVASIGNLQMDNATIAAALVYTKKEFQLQNEPTVE---GLIYMSLIGATITVTSBG 119
DB 26 AIVASIIISITFGYDVGWVGAGQIFLRDDIKI-NDQIEVLAIIMLCALVSGLT---AG 80
QY 120 PLSDSISGRPMILSSILYFSGILMLMGPNYVLLARFVDGFGIGLAVTLVPLYSRI 179
DB 81 KTSVDYIGRRYITIALANVIFVSGVLMGYPNYPVLMGICAGVGVGAPLMAPYSAEI 140
QY 180 APSEIRGLINTLTPQSGSGMPL---SYCWFGMSLSPSPDMRIMLGVALIPSLFF-FG 234
DB 141 SSASIRGFLTSIPELCISIGILGYVSNYC--FG-KLTCLKGRIMLGIAAPSLILAFG 197
QY 235 LTIFFLPSPRMLVSKGRMAEAKVYLQKLRGKDDVSGELSLLEGLFVGSDTISIEYIIG 294
DB 198 IT-TPSPSPRLVVGRLSEAKKIM----- 221
QY 295 PATEADDLVTDDGKEQITLYGPEBQSWIARPSKQIMLGSLASLRHSGMNVQSVPL 354
DB 222 --- 221
QY 355 MDPIYTLFGSVHENNPAQAGSMRSTLPFNFGSMFVSVDQAHKNEQMDENLHRDDEYAS 414
DB 222 ---VLVNTTEBEAER----- 234
QY 415 DGAGDYEDNLHSPILSRQATGAEKDIYHGHRSALSMRQITLLGEGDGVSTDIG 474
DB 235 ---FRDIL---TAAB---VDVTEIKEVGG 254
QY 475 GMDLMMKSEKENGKRGKGFYVYLHQGVPGSRGSIIVSLPGGDVFESEFVHAA 534
DB 255 G-----VKKKNKK----- 263
QY 535 LVSQSALFSGKLAERPMSDAAMVHPSEVAAKSRMKDLF--EPGVRALLVGVGIQLQ 591
DB 264 ---SVRELVIKRPAPVRIILLIAVGIHFE 291
QY 592 QPAGINGVLYYTPQILEQAVN---VLSKFGLSASASIISSITLILMPCIGPML 647
DB 292 HATGEAVLVSPRIFKAGVSKDLATVGVGLTKAFPII-----IATF 338
QY 648 LMDISGRFLLIGTIPILIASIIVIVSNLI--DIGTIAHAL-LSTVSIVYFCCFVWG 704
DB 339 LLDKGRKRLILTSVGVNFAITSLAVSLTWQGRGLAMALSLISIVTYAVAFSITGL 398
QY 705 GPIPIILCAEIPTRVRGLCIAIATFTWIGDIIVTYSIPVNLAIAGLAFSIVAVCL 764
DB 399 GPIITVYSEIFPLRLAQASISIGAVNRIMATVMSFSLMTKAITTGCVFPAIGAV 458
QY 765 ISFVFLVETKMPLEVITEFFAVG 792
DB 459 AAMWFFPMLPETKGLPLEMEKRLFGGG 486

```

RESULT 13

A85433
sugar transporter like protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001
C:Accession: A85433
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:2008348; PMID:10617198
A:Accession: A85433

A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-493 <STO>
A:Cross-References: GB:NC_001268; NID:g7270615; PIDN:CAB80333.1; GSPDB:GN00140
C:Genetics:
A:Gene: AT4g36670
A:Map position: 4
C:Superfamily: glucose transport protein

Query Match 12.1%; Score 493; DB 2; Length 493;
Best Local Similarity 23.2%; Pred. No. 7.1e-27;
Matches 177; Conservative 102; Mismatches 178; Indels 306; Gaps 23;

```

QY 48 SDDILEDKMSG---AVLVAIVASIGNLQMDNATIAAALVYTKKEFQLQNEPTVE--- 100
DB 2 ADQISGEPAVRNPAFLQALVAVISITFGYDVGWVGAGQIFLRDDIKI-NDQIEVL 60
QY 101 GLIYMSLIGATITVTSBGPLSDSISGRPMILSSILYFSGILMLMGPNYVLLARFV 160
DB 61 GILMLCALVSGLT---AGRTSDIIGRRYITVILASILFMLGSIIMGMPNYPVLLSGRCT 116
QY 161 DFGIGLAVTLVPLYSISIRGLINTLTPQSGSGMPLSCMFMSLSFSP-DWR 219
DB 117 AGLGVPALMVA PYSAEIAATASHRGLASLPHLCISIGILGYVSNYFSGKLPMTIGMR 176
QY 220 IMLGVLAIPSLFF-FGLTIFFLPSPRMLVSKGRMAEAKVYLQKLRGKDDVSGELSLLE 278
DB 177 LMLGIAVPSLVLAFFG--ILMPSPRMLVQGLKKEGKEL----- 216
QY 279 GLEVGDTSIEYIIGPATEADDLVTDDGKEQITLYGPEBQSWIARPSKQIMLGSLV 338
DB 217 ---ELVSN---SPEEAB----- 227
QY 339 SLARHSGMNVQSVPLMDPIYTLFGSVHENNPAQAGSMRSTLPFNFGSMFVSVDQAHKNE 398
DB 228 ---LR----- 229
QY 399 QMDENLHRDDEYASDAGDYEDNLHSPILSRQATGAEK--DIYHGHRSALSMR 455
DB 230 ---FQD-----ITAAAGIDPKCYDDV----- 248
QY 456 RQTLLEGDGVSTDIGGQMLAMKSEKENGKRGKGFYVYLHQGVPGSRGSIIV 515
DB 249 ---KMGKK-----TGEGV----- 260
QY 516 SLPGGDVFESEFVHAAALVQSALFSGKLAERPMSDAAMVHPSEVAAKSRMKDLF-- 573
DB 261 ---WKELILR 267
QY 574 -EPGVRALLVGVGIQLIQPAGINGVLYYTPQILEQAVNVLKFGLSASASIISS 632
DB 268 PTPAVRVLTLALGIIHFQHASGIEAVULYGPRIFKAGIITF-KDKLFLVITGIM--- 323
QY 633 LTTLLMPCIGPAMLLMDLGRFLLIGTIPILIASVITL-----VVSNLDIGTIAHAL 687
DB 324 KTFRIFT-----ATLLDKVGRKRLILTSVGVNFAITSLAVSLTWQGRGLAMAL 376
QY 688 -LSTVSIVYFCCFVWGPGPIPNILCAEIPTRVRGLCIAIATFTWIGDIIVTYSIPV 746
DB 377 VLSTVAAYSFAVAFSISIGIPITWYSSBVPPLKRLAQASISIGAVNRIMATVMSFSL 436
QY 747 LNAIGLAGVPSIVAVVCLISFVFLVETKMPLEVITEFF 789
DB 437 TSAITTGAFPMFPAVNAVAMNFFFLIPETKGLSEIEIALF 479

```

RESULT 14

G84864
probable membrane transporter [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: G84864
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanKen, S.E.; Unayam, L.; Tallon, L.; Euser, D.; Nierman, W.C.; White, O.; Eileen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J. Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A:Reference number: A04420; MUID:20083487; PMID:10617197

A:Accession: G84864

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-521 <STO>

A:Cross-references: GB:AB002093; NID:g2899003; PIDN:AA064332.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g43330

A:Map position: 2

C:Superfamily: glucose transport protein

Query Match 12.0%; Score 489; DB 2; Length 521;

Best Local Similarity 21.2%; Pred. No. 1.5e-26;

Matches 159; Conservative 98; Mismatches 179; Indels 314; Gaps 13;

62 VAIVASIGNLQGMNATIAAVALYIKKEFOLONEPTEGL-----IVSMSL 108

34 LTVNAGIGLFGVYTGVIISGLYIKODFEVVKSSFLQYNNVSSFTSSKLETTIVSMAL 93

109 IGATVITTFSGPLSDISGRPMILISILYFSGILIMSPNVYLLARFVDFGIGLA 168

94 VGAMIGAAGGWINDYGRKKTTLFADYFAAGATVMAAPDPYVLISGRLLVGLGVVA 153

169 VTLVVLVLTSEIAPSRIRGLNLTLPQSSGGMFLSYCVFMSLSPSPDMRMGLVLAIP 228

154 SVTAIVVYIAEASPSSEVRGLVSTNVMITGGQFLSYLVNSAFYVPG-TWRMVLGVSGVP 212

229 SLFFGLTIFLYPESPRMLVSKGRMAEAKVLOKRGKODVSGELSLLEGLVGGDSI 288

213 AVIQITLMLF-MPESPRLFMKAKKALQVLAAR----- 245

289 EBYIIGPATKADLVTDGKEQITLYGBEGQSIAPSPKPIMLGVLISLASHGSMV 348

246 ----- 245

349 NQSVLMPPIYTLRGSVHNNMQAGSMRSTLFPNFGSMFSTVDHAKNEQDENLRD 408

246 -----TYDISRL----- 252

409 DEEYASDAGDVEDNLHSPILSRQATGAEKDIYHGHRSALSMRQTLTGEGDGV 468

253 -----EDET-----DHLS 260

469 STDIGGQWLAMKMSKEGNGRKEGFRVYVHOGVPSRRGSIIVSLPGGDVFESE 528

261 AAE-----EBEKQRKRTVG--YL-----DVFRSKE 283

539 FVHAALVYQSALBEKGLAEPRMSDAMVHSEVAAKSRWKDLFEPGRALLVGVGIQ 588

284 -----LRLAFLAGAGIQ 295

589 ILQFAGINGVLYTTPOLLEAG-----VAVILSKRGLSSASATLISLTLLMLPCIG 643

296 AFQQTGINTVMTYPTIYQNAHFSNQALFLSLIVAMNAAGTV----- 342

644 PAMLLMDISGRRLILGTIPILIASLVILVSNLIDLGLTA-----HALSTSVIIVFCC 699

343 -GIYIDHGRKKLALSLFVITISLILSVSFFQSESTSDGGLYGMALVGLALYIVF 401

700 FVMGEGPIPNILCAIFPFRVAGLCAICAFTFWIGDIIIVTYSLPVMLAIGLAVFSIY 759

402 FAPGEGPVPWTVNSIYPOOYRGICGMSATNMISNLIVAOQTFILIAEAAGTGMTFLIL 461

760 AVVCLISPFVFLKVPETKGMPLVETFEF 789

462 AGIIVAVLIVLVFVPEIQLTFSEVEQIV 491

T14606
probable sugar transport protein 205 - beet

C:Species: Beta vulgaris (beet)

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 01-Dec-2000

C:Accession: T14606; T14617

R:Chou, T.J.; Bush, D.R.

Plant Physiol. 110, 511-520, 1996

A:Title: Molecular cloning, immunochemical localization to the vacuole, and expression

A:Reference number: Z18131; MUID:96351183; PMID:8742332

A:Accession: T14606

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-549 <CHI>

A:Cross-references: EMBL:U64902; NID:g1778092; PID:g1778093

A:Accession: T14617

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-541, 'SVQV' <CH2>

A:Cross-references: EMBL:U64903; NID:g1778094; PID:g1778095

C:Genetics:

A:Note: Bvcdna-205; Bvcdna-397

C:Superfamily: glucose transport protein

C:Keywords: sugar transport

Query Match 12.0%; Score 488; DB 2; Length 549;

Best Local Similarity 21.3%; Pred. No. 1.8e-26;

Matches 164; Conservative 113; Mismatches 194; Indels 298; Gaps 19;

41 SQSEPTVSDILED-----KMSGAVLVAIVASIGNLQGMNATIAAVALYIKKEF 91

10 SDPPTTASKVIADFDPLKPKPKKPKKFAFACATLASMTSVILGIDIGMSAGIITLKEDW 69

92 QLONEPTVEGLIVMSLIGATVITTFSGPLSDISGRPMILISILYFSGILIMSPNV 151

70 HISD--TQIGLVGLINLYCLFGSPAAGRTSDWIGRRITVLAGAIFVGLALMGFAINY 127

152 YVLLARFVDFGIGLAVTLVPLIYSEIAPSRIRGLNLTLPQSSGGMFLSYCVFMS 211

128 AFLWGRVITIGVYALMIAVPYTAESPPASSGFLSPFEVFINIGILIGYISNLAFS 187

212 LSPSP-DWRIMGLVLAIPSLFFGLTIFLYPESPRMLVSKGRMAEAKVLOKRGKDDVS 270

188 SLPTLSWRFMIGICALPSI-FLAIGVLAEPSPRLVMQRLDPAKKVLNR----- 238

271 GELSLLEGLVGGDTSIEEYIIGPATKADLVTDGKEQITLYGBEGQSIAPSPK 330

239 -----ISD-----SPEAQ----- 247

331 PIMLGSVLSLRHGMVNSQVPLMDPIYTLFGSVHNNMQAGSMRSTLFPNFGSMFSV 390

248 ----- 247

391 TDQAKNEQWDEENLRHDEEYASDAGDVEDNLHSPILSRQATGAEKDIYHGHRS 450

248 -----L 248

451 ALSMRQTLTGEGDGVSTDIGGQWLAMKMSKEGNGRKEGFRVYVHOGVPSRR 510

249 RLSEIKQT-----AIP- 261

511 RGSIVSLPGGDVFESEFVAALVQSALFSKGLAEPRMSDAMVHSEVAAGSRWK 570

262 -----ECDE-----DIYKVKTKIKSNAVWK 283

571 DLFF--EBGVRALLVGVGIQILOQFAGINGVLYTTPOLLEAGVAVILSKRGLSSASAS 627

284 ELFENPTAVRRVAVIAGIGIFFOQASGIDAVVIVSPRIFQSAGITNARKQL-----LAT 338

628 ILISLTLLMLPCIGFPMMLMDISGRRLILGTIPILIASLVILVSNLIDLGLTA-----T 682

339 VAVGVVKTFLTL-----VATFQLDKYGRRPLITVSGVMIALTLTLMSTVYIDSHHKIT 394

683 LAHALSTSVIIVFCCFVMGFGPIPNILCAIFPFRV--GLCAICAFTFWIGDIIIVT 740

Db 395 WAIALCITM-YCAVVASFSIGLGPITWYSSSEVFLRLRAOGTSMGV-AVNRVSGVTSI 452

OY 741 YSLPVMANAGIAGVPSIYAVVCLISFVFLKVPETKGMPLVITEFF 789

Db 453 FFLP-LSHKITTGAFPLFGIAIIAMFFFLTLPETRGRTLENMHELP 500

Search completed: January 5, 2004, 18:58:44
Job time : 49 secg

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 15:50:52 ; Search time 24 Seconds
(without alignments)
1567.557 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGSWLAVQFPFPDLDRL.....PLEVITPEFAVGAQAAAKA 800

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	527	12.9	461	1	CSBC_BACSU
2	486.5	11.9	482	1	YFBC_BACSU
3	483	11.9	457	1	XYLT_LACBR
4	473.5	11.6	464	1	ARAE_BACSU
5	462	11.3	612	1	ITR2_YEAST
6	451	11.1	464	1	GAIP_ECOLI
7	434.5	10.7	472	1	ARAE_ECOLI
8	434	10.6	557	1	ITR2_SCHPO
9	432	10.6	547	1	GTR1_LEIDO
10	431	10.6	547	1	ARAE_ECOLI
11	429.5	10.5	472	1	ARAE_ECOLI
12	421.5	10.3	584	1	ITR1_YEAST
13	421	10.3	629	1	MYCT_HUMAN
14	418.5	10.3	468	1	GLCP_SYNY3
15	404.5	9.9	575	1	ITR1_SCHPO
16	384	9.4	533	1	GTR2_CHICK
17	382	9.4	534	1	HDP3_CHICK
18	364.5	8.9	570	1	HXT1_YEAST
19	354	8.7	492	1	GTR1_BOVIN
20	352.5	8.7	541	1	HXT2_YEAST
21	352	8.6	492	1	GTR1_MOUSE
22	352	8.6	522	1	STR_RICCO
23	350.5	8.6	478	1	GTR8_RAT
24	350	8.6	477	1	GTR8_MOUSE
25	350	8.6	492	1	GTR1_RAT
26	349.5	8.6	592	1	HXT5_YEAST
27	348.5	8.6	522	1	GTR2_RAT
28	348.5	8.6	523	1	GTR2_MOUSE
29	348	8.5	492	1	GTR1_HUMAN
30	346.5	8.5	534	1	HDP1_CHICK
31	346.5	8.5	574	1	GAU2_YEAST
32	346	8.5	524	1	GTR2_HUMAN
33	345.5	8.5	546	1	GHT5_SCHPO

34	345	8.5	490	1	GTR1_CHICK	P46896 gallus galli
35	344	8.4	509	1	GTR4_HUMAN	P14672 homo sapien
36	343	8.4	477	1	GTR8_HUMAN	Q9ny64 homo sapien
37	343	8.4	566	1	KHT2_KLITA	P53387 kluyveromyc
38	343	8.4	567	1	HXT3_YEAST	P32466 saccharomyc
39	342	8.4	496	1	GTR1_RABIT	P13355 oryctolagus
40	342	8.4	496	1	GTR3_CHICK	P28568 gallus galli
41	342	8.4	570	1	HXT6_YEAST	P39004 saccharomyc
42	342	8.4	570	1	HXT7_YEAST	P39003 saccharomyc
43	341	8.4	576	1	HXT4_YEAST	P32467 saccharomyc
44	339	8.3	507	1	GTR6_HUMAN	Q9ugq3 homo sapien
45	339	8.3	522	1	STR1_ARATH	P23586 arabidopsis

ALIGNMENTS

RESULT 1
CSBC_BACSU STANDARD; PRT: 461 AA.
ID CSBC_BACSU
AC P46333; Q32289; (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable metabolite transport protein csbc.
GN CSBC OR S8928R.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / BGSC1A1;
RX MEDLINE=96093926; PubMed=7584049;
RA Yoshida K.-I., Seki S., Fujimura M., Miwa Y., Fujita Y.;
RT "Cloning and sequencing of a 36-kb region of the Bacillus subtilis
RL genome between the gnt and iol operons.";
RN DNA Res. 2:61-69(1995).
[2]
RP REVISIONS.
RA Fujita Y., Shibayama T., Ishio I., Aoyama D., Yoshida K.-I.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
[3]
RC SEQUENCE FROM N.A.
RX STRAIN=168;
RA MEDLINE=96044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bortner M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriss R., Boursier L., Brans A., Braun M., Brigelli S.C., Bron S.,
RA Broillet S., Brunsch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denton F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson F.T.,
RA Euteneier K.D., Ewington J., Fabret C., Ferrati E., Foulger D.,
RA Fritz C.Y., Glaeser P., Goffeau A., Golligly E.J., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golligly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holtsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kaashara Y., Klaetr-Blanchard M., Klein C.,
RA Kobayashi Y., Koeltter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Laber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogawara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik G., Prescott A.M.,
RA Priesen E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha B., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,
RA Sekiguchi J., Sekowaka A., Seror S.J., Serro P., Shin B.S., Soldo B.,
RA Sorokin A., Taccioni E., Takagi T., Takahashi H., Takenaru K.,
RA Takuchi M., Tamakoshi A., Tanaka T., Terstera P., Tognoni K.,
RA Toesto V., Uchiyama R., Vandenbol M., Vannier F., Vassaret A.,
RA Viati A., Wambut R., Wedler B., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;

RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 subsp. ";
 RT Nature 390:249-256(1997).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=99303315; PubMed=10376822;
 RA Akbar S., Lee S.Y., Boylan S.A., Price C.W.;
 RT "Two genes from *Bacillus* subsp. under the sole control of the
 RT general stress transcription factor sigmaB.";
 RL Microbiology 145:1069-1078(1999).
 CC -1- FUNCTION: COULD SERVE EITHER A NUTRITIONAL OR AN OSMOTIC
 CC PROTECTION FUNCTION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AB005554; BAA21604.1; -
 DR EMBL; Z99124; CAB16017.1; -
 DR PIR; D70073; D70073.
 DR Subtilist; BG11360; cshC.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; SUG_TRANSPORTER.
 DR InterPro; IPR005829; SUG_TRANSPORTER.
 DR InterPro; IPR003663; SUGAR_TRANSP.
 DR Pfam; PF00083; sugar tr; 1.
 DR PRINTS; PR00171; SUGTRANSPORT.
 DR TIGR; TIGR00879; SP; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transport; Transmembrane; Complete proteome.
 FT DOMAIN 1 14
 FT TRANSMEM 15 35
 FT DOMAIN 36 38
 FT TRANSMEM 39 59
 FT DOMAIN 60 76
 FT TRANSMEM 77 97
 FT DOMAIN 98 104
 FT TRANSMEM 105 125
 FT DOMAIN 126 139
 FT TRANSMEM 140 160
 FT DOMAIN 161 163
 FT TRANSMEM 164 184
 FT DOMAIN 185 241
 FT TRANSMEM 242 262
 FT DOMAIN 263 280
 FT TRANSMEM 281 301
 FT DOMAIN 302 308
 FT TRANSMEM 309 329
 FT DOMAIN 330 341
 FT TRANSMEM 342 362
 FT DOMAIN 363 378
 FT TRANSMEM 379 399
 FT DOMAIN 400 402
 FT TRANSMEM 403 423
 FT DOMAIN 424 461
 FT CONFLICT 400 401
 SQ SEQUENCE 461 AA; 50235 MW; B6587B5DC7272EF1 CRC64;
 Query Match 12.9%; Score 527; DB 1; Length 461;
 Best Local Similarity 21.5%; Pred. No. 6.4e-27;
 Matches 156; Conservative 101; Mismatches 170; Indels 300; Gaps 12;
 QY 61 LVAVASIGNLQGDNAITAAVYIKKEFQLEPEVEGLIVMSLIGATITTFSGP 120
 DB 9 MIYFGALGGLGYDTGVISGALLFINNDIPLTT--LLEGIVVMLLIGAFSGALSGT 66

QY 121 LQSGRRPMLISLIVFFSGLIMLSPNVYLLARFVDFGIGLAVTLVPIYSEIA 180
 DB 67 CSDRMGRKRVFVLSIIFFIGALACAFSGITGMILASRVILGLAVGSGTALVPIYLSMA 126
 QY 181 PSEIRGLINTLPQFSGGGMFLSYCMVFGMSLSPSPWRIMLVGLAIPSLFFGLITPYL 240
 DB 127 PTKIRGTGTGNNMIMIVGILLAY--IVNYLFTPEAMRWVGLAVAVALL--IGIAFM 183
 QY 241 PESPRLVSKRMMAKRVQKLRGKODVSGELSLLEGLVGGDTTSEIYIIPATFA 300
 DB 184 PESPRLVSKRSEEDARIMKITHDPKDIENELMKNQG----- 222
 QY 301 DDLVDPGKEQITLYGPEEGSWIARPSKPIMLGSLASRRGSMVNGSPIMLPIVT 360
 DB 223 -EAKKETT----- 231
 QY 361 LFGSVHNMPOAGSMSTLFPNFGSMFVTDQAKNEQWDENLHRDDEYASDAGCD 420
 DB 232 ----- 231
 QY 421 YEDNLHSLRSQATGAEGKDIYVHGHRSALSMRROTLLGEGGDVYSTDIGGWQLAW 480
 DB 232 -----LKA 236
 QY 481 KMEKEGNGRKEGFKRVYLHQGVPGSRGIVSLPGGDVFESEFVAALVQSQA 540
 DB 237 KM----- 238
 QY 541 LPSKGLAPRMSDAMVHPSEVAAKSRWKDLFEPGVRRALLVGVGIQIIOQFAGINGVL 600
 DB 239 -----IRPMLIIVGLAIFQAAVGINTVI 262
 QY 601 YTPQILEQAVNIIKFGLSASASATLISLTLLMPCITGRAMIMLDSGRRLTL- 659
 DB 263 YVAP-----TIFTKAGLSTASALGTWGIILNVIKTI-TAMILIDVGRKLLTW 312
 QY 660 GTIPILLASVILVSMILIDGTLAH-ALSTVSVIYFCFCFWMGFGPIPIILCAEIPPT 718
 DB 313 GSVGTTLSLALSGV--LITGLSASTAMKIVPLGVIYVQATWGVVWVLMPELPS 370
 QY 719 RVRLGICAIKFTWIGDIIYVTSLPVNLNAGIAGVSIYAVVCLISFVFLKVPETK 778
 DB 371 KARAAVNGFTLLVLSAANLIVSLVFLPMLRPMGIAMVFMFVSICLISFFAFYVWPETK 430
 QY 779 GMPLEVI 785
 DB 431 GKSLSEI 437
 RESULT 2
 YFIG_BACSU STANDARD; PRT; 482 AA.
 AC P54723;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical metabolite transport protein yfig.
 GN YFIG.
 OS *Bacillus* subsp.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.
 OX NCBI_Taxid=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=96262713; PubMed=8704981;
 RA Yamamoto H., Uchiyama S., Fajar A.N., Sekiguchi J.;
 RT "Determination of a 12 kb nucleotide sequence around the 76 degrees
 RT region of the *Bacillus* subsp. chromosome.";
 RL Microbiology 142:1417-1421(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;

RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borcherst S.,
 RA Borries R., Bourrier L., Brans A., Braum M., Brigelli S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Denicof F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrati E., Foulger D.,
 RA Filtz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
 RA Giuseppe G., Guy B.J., Haga K., Haech U., Harwood C.R., Henauf A.,
 RA Hilbert H., Holtsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Katamata D., Kasahara Y., Klaert-Blanchard M., Klein C.,
 RA Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,
 RA Kunita K., Lapins A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
 RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeder R., Scoffone F.,
 RA Setiguchi J., Sekowaka A., Serron S.J., Serron P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi B., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpsira P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambuit R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zimstein E., Yoshikawa H., Zanchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 subtilis";
 RL Nature 390:249-256(1997).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL; D50543; BAA09111.1; -
 DR EMBL; Z99108; CAB12655.1; -
 DR PIR; B69803; B69803.
 DR Subtilist; BGI1854; YFIG.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub_transporter.
 DR InterPro; IPR005829; Sub_transporter.
 DR InterPro; IPR003663; Sugar_transpr.
 DR Pfam; PF00083; sugar_tr.1.
 DR PRINTS; PR00171; SUGERTNSPORT.
 DR TIGRFAMS; TIGR00879; SP.1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; FALSE_NEG.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Hypothetical protein, Transport, Transmembrane, Complete proteome.
 FT DOMAIN 1 29
 FT TRANSMEM 1 29
 FT DOMAIN 30 50
 FT TRANSMEM 30 50
 FT DOMAIN 51 59
 FT TRANSMEM 51 59
 FT DOMAIN 60 80
 FT TRANSMEM 60 80
 FT DOMAIN 81 92
 FT TRANSMEM 81 92
 FT DOMAIN 93 113
 FT TRANSMEM 93 113
 FT DOMAIN 114 120
 FT TRANSMEM 114 120
 FT DOMAIN 121 141
 FT TRANSMEM 121 141
 FT DOMAIN 142 155
 FT TRANSMEM 142 155
 FT DOMAIN 156 176
 FT TRANSMEM 156 176
 FT DOMAIN 177 184
 FT TRANSMEM 177 184
 FT DOMAIN 185 205
 FT TRANSMEM 185 205
 FT DOMAIN 206 263
 FT TRANSMEM 206 263
 FT DOMAIN 264 284
 FT TRANSMEM 264 284
 FT DOMAIN 285 301
 FT TRANSMEM 285 301
 FT DOMAIN 302 322
 FT TRANSMEM 302 322
 FT DOMAIN 323 331
 FT TRANSMEM 323 331

FT TRANSMEM 332 352 9 (POTENTIAL).
 FT TRANSMEM 353 373 10 (POTENTIAL).
 FT DOMAIN 374 400 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 401 421 11 (POTENTIAL).
 FT DOMAIN 422 443 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 424 444 12 (POTENTIAL).
 FT DOMAIN 445 482 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 446 482 13 (POTENTIAL).
 SQ SEQUENCE 482 AA; 52756 MW; 90851C4F4C48EED01 CRC64;
 Query Match 11.9%; Score 486.5; DB 1; Length 482;
 Best Local Similarity 20.7%; Pred. No. 2.96-24;
 Matches 158; Conservative 109; Mismatches 195; Indels 303; Gaps 15;
 39 SCSSQEPYTDSDIEDKMGSGAV-LVAVASGNTLQGMNATIAAVLYIKKEPOLNEP 97
 2 STKKKEAVTGKESLAHK--GLRTITLVSTFGGLPGYDGVNGLALPMAATAGQLNUTP 59
 98 TVEGLIVMSPLIGATITVTTSGPDSIGRRPMLISIIYFSGLIMLSPNYVLLLA 157
 60 VTEGLIVASSLLGAFGAFMGGRSDHGRKRTILYALFLPFIATLCTGTPSNASVMIAT 119
 158 RFVDGFGIGLAVTVLPYIETIASETRGLINTLPQSGSGCMFLSCM--VFQMSLSPS 215
 120 RFLGLAVGCASVTVPTFLAETSPERRGRIVTQNEIMIVIGQLAVTFNAIIGSTWGES 179
 216 PD-WRMIVGLVAIPS-LFEGELTFYLPESPRLVSKGRMAKKVLOKRGKDVSGEL 273
 180 ANVRRVLLVATLVAVALVWFQML--VPESRWLAAGRGMDALRVLRQIR----- 228
 274 SLLEGLVGGDTIEEYIIGPATEAADLVTDGKEQITLYGEEGSGWIARPSKPIM 333
 229 -----EDGQAQ----- 235
 334 LGSVLASLRHGSVNGSVPLMDPIVTLFGSVHNMPOAGSMSTLFPNFGMSFVTDQ 393
 236 ----- 235
 394 HAKRGQDEEVLHRDDEYASDAGGDEYDNLHPSLSRQTGAEGKDIYHGRGSALS 453
 236 -----EIKETIKI----- 242
 454 MRQTLLEGDGDVSTDIGGQWLAWKSEKEGNGRKEGFRVYVHQGVGSRGS 513
 243 ----- 242
 514 IVSLPGGDVEEGSEFVHAALVYSQALFSKGLAEPKMSDAVMPSEVAKGSRWDLF 573
 243 -----ALEGT-----AKKGFHDFQ 257
 574 EPGVRALLVVGQIIQOPFNGINGVLYTTPQIIIEQKGVAVILSKFGLSASASALISL 633
 258 EPMIRILFIDIGIAYQOITGVNSIYVGTIELIREAGF-----QTEALIGIA 307
 634 TTLMLPFCIGPMLMDISGR-FLKGTPIILASVILVNSMLDGTALHALLSTVS 692
 308 NGVSVLAVITGVMLGKVRRRPMLIIGQITMALHIGLISVLE-CTPALRYV-VLS 365
 693 VIYVFCFVM-GFSPINILCAEIPTRVRGLCAICAFTEWIGDIIIVYSLPVMAIG 751
 366 LTIPLFAQCAISTVWMLMSIFPMHVRGLGMSIFFLCMTANFLIGFPILLNHIG 425
 752 LAGPSIYAVVCLISFPVPLKVPETKMPLEVTETPEPAVAKQA 796
 426 MSATFFIFVANNILAILFVKKYVETKGRSLQLEHSPROYGRA 470
 RESULT 3
 XYLUT_LACBR STANDARD; PRT; 457 AA.
 ID XYLUT_LACBR
 AC 052733;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)

RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
RA Paro T.M., Portetle D., Portetle D., Prescott A.M.,
RA Priesen E., Puig C., Purnelle B., Rapoport G., Ray M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan B., Schleich S., Schroeter R., Scoffone F.,
RA Setoguchi J., Sekowska A., Seror S.J., Serron P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Yamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier G., Vassart A.,
RA Viati A., Wambuit R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
RT *subtilis*.";
RT Nature 390:249-256(1997).
RL [2]
RN SEQUENCE OF 1-223 FROM N.A.
RP STRAIN=168;
RC MEDLINE=97197523; PubMed=9045819;
RX Sa-Nogueira S., Mota L.J.;
RA "Negative regulation of L-arabinose metabolism in *Bacillus subtilis*:"
RT characterization of the *araC* gene.";
RL J. Bacteriol. 179:1598-1608(1997).
RN [3]
RP SEQUENCE OF 223-464 FROM N.A., AND FUNCTION.
RC STRAIN=168;
RX MEDLINE=98062200; PubMed=9401028;
RA Sa-Nogueira I., Ramos S.S.;
RT "Cloning, functional analysis, and transcriptional regulation of the
RT *Bacillus subtilis* *araC* gene involved in L-arabinose utilization.";
RL J. Bacteriol. 179:7705-7711(1997).
RN [4]
RP TRANSCRIPTIONAL REGULATION.
RC MEDLINE=99348379; PubMed=10417639;
RX Mota L.J., Tavares P., Sa-Nogueira I.M.G.;
RT "Mode of action of *AraC*, the key regulator of L-arabinose metabolism
RT in *Bacillus subtilis*.";
RL Mol. Microbiol. 33:476-489(1999).
RN -1-
RP FUNCTION: Uptake of arabinose across the boundary membrane with
RC the concomitant export of a proton (symport system).
RX -1-
RT SUBCELLULAR LOCATION: Integral membrane protein.
CC -1-
CC INDUCTION: Transcription is repressed by the binding of *araC* to
CC the promoter. L-arabinose acts as an inducer by inhibiting the
CC binding of *araC* to the DNA, thus allowing expression of the gene.
CC -1-
CC SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; 299121; CAB15401.1; -
CC EMBL; X98354; CA66998.1; -
CC EMBL; Y12105; CAAT2812.1; -
CC PIR; F69587; F69587;
CC Subtilist; BG11907; *araC*.
CC InterPro; IPR007114; MFS.
CC InterPro; IPR005828; Sub_transporter.
CC InterPro; IPR005829; Sug_transporter.
CC InterPro; IPR003663; Sugar_transp.
CC Pfam; PF00083; sugar_tr; 1_
CC PRINTS; PR00171; SUGTRANSPORT.
CC TIGRfam; TIGR00879; SP; 1_
CC PROSITE; PS00216; SUGAR_TRANSPORT_1; 2_
CC PROSITE; PS00217; SUGAR_TRANSPORT_2; 1_
CC Transport; Sugar transport; Transmembrane; Symport; Complete proteome.

FT TRANSMEM 21 43 POTENTIAL.
FT TRANSMEM 63 85 POTENTIAL.
FT TRANSMEM 92 111 POTENTIAL.
FT TRANSMEM 116 138 POTENTIAL.
FT TRANSMEM 150 172 POTENTIAL.
FT TRANSMEM 185 207 POTENTIAL.
FT TRANSMEM 266 288 POTENTIAL.
FT TRANSMEM 303 325 POTENTIAL.
FT TRANSMEM 332 354 POTENTIAL.
FT TRANSMEM 364 386 POTENTIAL.
FT TRANSMEM 398 420 POTENTIAL.
FT TRANSMEM 424 446 POTENTIAL.
SQ SEQUENCE 464 AA; 50411 MM; 13B417061CB61DA2 CRC64;
Query Match 11.6%; Score 473.5; DB 1; Length 464;
Best Local Similarity 20.0%; Pred. No. 1.9e-23;
Matches 153; Conservative 102; Mismatches 191; Indels 319; Gaps 13;
QY 37 PASCSQEPVTSDDLLBDMKSGAVLVAIVASIGNLQGMNATTAALVLYIKKEFOLONE 96
DB 5 PQLLEPNVPVTR-----SHSMGFVLISCAAGLGLVGYDVAIVISALGFLKDLVSL--S 58
QY 97 PTEGLIVSMGLIGATVITTSGLPSDSIGRPMTILSIIFYPSGLIMNSPNVYVLL 156
DB 59 PEMEGLIVSIMGVGVGVSIGFSDRGRKIMTALFAISAIVSALSQDVSTLII 118
QY 157 ARFVDFGIGLAVTLVPIYISEIASEIRGLINTLPQSSGGMFLSCMFGMSLSPSP 216
DB 119 AKTIGGLIGMSSLSVTYTEAPPAIRGLSSLSYQFTLLIGSATYFINLAVRSSTY 178
QY 217 D-----NRIMGLVLAIPSLFFPGITLYLPSPFMTLSKSRMAEAKVYLQLRKQDVS 270
DB 179 EMGVTGRKMLAYGMVSVIFF-LVLLVESPRLAKAKTNEALILIRINSEYAK 237
QY 271 GELSLLEGLVEGDTSTIEYIIGPTEADDLVTDGKEQITLYGPREGSWIARPSKG 330
DB 238 ELKNIENSLKI-----EQM----- 252
QY 331 PIMGLSVSLASRHSWMNQSVPLMDPIVTLFGSVHEMNPQAGSMRSTLPFGSMFSV 390
DB 253 ----- 252
QY 391 TDQAKNEQWDBENLRDDEEYASDAGDYEDNLHSLRQATGAEKDIVHHGRGS 450
DB 253 ----- 252
QY 451 ALSMRQTLLEGGDVSTDIGGMQAMKSEKENGKRGKRVYLHQEGVPSR 510
DB 253 ----- 252
QY 511 RGSIVSLGGGVDFEGSEFVAHALVSQALFSGLAEPFRSDAMHPSEVAAGSRWK 570
DB 253 -GSL-----S 256
QY 571 DLFEGVRRALLVGVIQILOPAGINGVLYTTPQILEQAGAVAILSKRGLSASASILI 630
DB 257 QLFKPKKALVIGLILFNVQIGMAIITYGPPIFCMCG-----FGQAGVTTCTI 309
QY 631 SLLTILMLPCIGFAMLLMDSGRFL-LGTIPLIASVLVLSNLDLGT-----LA 684
DB 310 VGVEVAVIFTV--LAVLLIDKVRKMSIGS--AFMAIPMLI-----QTSYFPELT 357
QY 685 HALSTVSVIYVFCFVNGFRGIPNILCAEIFPTTVRGLCIAICFTWIGDIITYSLP 744
DB 358 SGIMMIIVILGVAAFCVSVGPITWIMSEIFPNILARRAAGIATIFLMGAMWALGOFPV 417
QY 745 WLNALIGAGVSIYAVCLISFVFLKVPETKMPLEVTIEFF 789
DB 418 MMIDSGLAITYTWIYAVINILCLFVVTICETKKSLEIEKLM 462
RESULT 5
ITR2_YEAST

ID	ENTRY	STANDARD:	PRT:	612 AA.
DT	01-APR-1993	(Rel. 25, Created)		
DT	01-APR-1993	(Rel. 25, Last sequence update)		
DT	28-FEB-2003	(Rel. 41, Last annotation update)		
DE	Myo-inositol transporter 2.			
GN	ITR2 OR YOL103W OR HRB612.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.			
OX	NCBI_TaxID=4932;			
RA	SEQUENCE FROM N.A.			
RA	MEDLINE=91250431; PubMed=2040626;			
RT	Nikawa J.-I., Tsukagoshi Y., Yamashita S.;			
RT	"Isolation and characterization of two distinct myo-inositol			
RT	transporter genes of Saccharomyces cerevisiae.",			
RL	J. Biol. Chem. 266:11184-11191(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96076631; PubMed=7502582;			
RA	Vandenbol M., Durand P., Portetelle D., Hilger F.;			
RT	"Sequence analysis of a 44 kb DNA fragment of yeast chromosome XV			
RT	including the Tyl-H3 retrotransposon, the sufl(+) frameshift			
RT	suppressor gene for crNA-Gly, the yeast transfer RNA-Thr-1a and a			
RT	delta element.";			
RL	Yeast 11:1069-1075(1995).			
CC	- FUNCTION: MINOR TRANSPORTER FOR MYO-INOSITOL.			
CC	- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sdb.ch/announce/			
CC	or send an email to license@isb-sdb.ch).			
CC	-----			
DR	EMBL; D90353; BAA14367.1; -.			
DR	EMBL; Z48145; CAA88159.1; -.			
DR	EMBL; Z74845; CAA99119.1; -.			
DR	PIR; B40538; B40538.			
DR	SGD; S0005463; ITR2.			
DR	GO; GO:0005365; F:myo-inositol transporter activity; IMP.			
DR	GO; GO:0015798; P:myo-inositol transport; IMP.			
DR	InterPro; IPR007114; MFS.			
DR	InterPro; IPR005828; Sub_transporter.			
DR	InterPro; IPR005829; Sug_transporter.			
DR	InterPro; IPR003663; Sugar_transp.			
DR	Pfam; PF000083; sugar_tr.1.			
DR	TIGRFAMS; TIGR00879; SP.1.			
DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.			
DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.			
RW	Transmembrane; Sugar transporter; Glycoprotein.			
FT	DOMAIN 1	107		
FT	TRANSMEM	108	128	
FT	DOMAIN	129	155	
FT	TRANSMEM	156	176	
FT	DOMAIN	177	182	
FT	TRANSMEM	183	203	
FT	DOMAIN	204	212	
FT	TRANSMEM	213	233	
FT	DOMAIN	234	241	
FT	TRANSMEM	242	262	
FT	DOMAIN	263	272	
FT	TRANSMEM	273	293	
FT	DOMAIN	294	362	
FT	TRANSMEM	363	383	
FT	DOMAIN	384	402	
FT	TRANSMEM	403	423	
FT	DOMAIN	424	426	
FT	TRANSMEM	427	447	
FT	DOMAIN	447		
FT	TRANSMEM	447		

Query Match	11.3%; Score 462; DB 1; Length 612;	
Best Local Similarity	19.9%; Pred. No. 1.5e-22;	
Matches 161; Conservative 112; Mismatches 202; Indels 334; Gaps 19		
DOMAIN 448 467	EXTRACELLULAR (POTENTIAL).	
FT TRANSMEM 468 488	10 (POTENTIAL).	
FT DOMAIN 489 512	CYTOPLASMIC (POTENTIAL).	
FT TRANSMEM 513 533	11 (POTENTIAL).	
FT TRANSMEM 534 536	EXTRACELLULAR (POTENTIAL).	
FT TRANSMEM 537 557	12 (POTENTIAL).	
FT DOMAIN 558 612	CYTOPLASMIC (POTENTIAL).	
FT CARBOHYD 397 397	N-LINKED (GLICNAC. . .) (POTENTIAL).	
SQ SEQUENCE 612 AA; 67041 MW; 48E733102BAD7298 CRC64;		
Query Match	11.3%; Score 462; DB 1; Length 612;	
Best Local Similarity	19.9%; Pred. No. 1.5e-22;	
Matches 161; Conservative 112; Mismatches 202; Indels 334; Gaps 19		
37 PACSSQEPVTSDDTLIEDKMSGAVLV-----ATVASCIN 70		
Db PVSTSTWQIKSRQD--EDEDGRIVIKFVNEDDPTSVITFNOSISPTLTLPVASISG 123		
QY 71 LLQGDNATIAAVALVIKKEPQLQNEPVEG---LIYMSLIGATIVTTFGSPLSDSIGR 127		
Db 124 FMFGYDTGYISSALISINRD--LNKRVLTYSKEKLTITATSIGALITVGAQTAADVGR 181		
QY 128 RPLMLISLIFFGSGILMLMSPNVVLILARFVDFGIGLAWTLVPLVYSIETAPSEIRGL 187		
Db 182 RPLCMFSMLMELIGAILDITAHKKPQMAAGLIMFGVIGISLISPLTISLAPKIRK 241		
QY 188 LNTLPQFSGSGMELSYCMVFMSLSPSPDRIMLVGLAIPSLFPGGLTFIYLPSPRWL 247		
Db 242 LTVINSLMLTGGQLIAYCCAGLNLHVANGWRILIVGLSLIPVLQFSPFCF-LPDTPRY 299		
QY 248 VSKGRMAAKVYLQKLRKDKDVSGELSLLEGLBEVGDTSLSEYIIIGATEAADLVYDG 307		
Db 300 VMKQDPLKAKKAVLKR-----SYV-----NTEDEIIDQ 326		
QY 308 DKEQITLYGPEGGSWILARPSKGPIMLGSVYSLASRHSQMVNOSVPLMDPTVLTGSHYE 367		
Db 327 KYEEL-----SLANGSLPGKNPITKFMNMYKE 353		
QY 368 NMPQAGSMRSTLPPNFGSMFSTVDQNAKNEQWDENLHRDEBYASDAGAGDYEDNLHS 427		
Db 354 -----LHT 356		
QY 428 PLLSRQATGAGCKDVIHHGHRGSALSMRKQTLLEGSGDGVSTDIGGQWLAMKMSKEG 487		
Db 357 ----- 356		
QY 488 ENGRKEGGFKRVYLHQEGVPSRRSGISYSLPGCGDVFEGSEFVHAALVSGSLPSKGLA 547		
Db 357 -----VP-----SNP----- 361		
QY 548 EPRMSDAAVHPSEVAAKGSRMKDLFEQGVRRALLVGVGILOQFAGINGVLYTTPQIL 607		
Db 362 -----RALIIGGGLAQIQFTGMNLSMTVSGTIF 390		
QY 608 EQAGVAVILSKFGLSSASASILISSLTTLMLPCIGFAMLMIDLSGRFPLIGTIPILIA 667		
Db 391 ETVGFK-----NSSAVSIIIVSGTNVEFTL---IAFCIDIKIGRRYIILIGIPGMTV 438		
QY 668 SLVILVVS-----NLIDL-----GLTAAALLSTVSVIYVPCFNWNGPPIPIILCAE 714		
Db 439 ALVICAIAFHFLLGIKFNAGADVVASDGSMSGVIYIVETIYYAAFYALGIGIVP-WQOSE 497		
QY 715 IFPRVRGGLCAIACAFTFMIDIIIVTYSLPMLNAIGLAGVFSIYAIVCLISFVFPFLKV 774		
Db 498 LFPQNVKGVGYSYATATNMAGSLVYASFLTMLQNTITPTGTFSPFAGVACLSITFCYCY 557		
QY 775 PETKGMPL-----VITEFFAVGAKQAOKA 799		
Db 558 PELSGLELEVEQVITLKQGFNIKASKALAK 586		

ID GALP_ECOLI STANDARD; PRT; 464 AA.
AC P37021;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Galactose-proton symporter (Galactose transporter).
GN GALP OR B2943 OR C3529.
OS Escherichia coli, and
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI TaxID=562, 217992;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RA Roberts P.E.;
RA Theis (1992), University of Cambridge, U.K.
RL [2].
RN SEQUENCE FROM N.A.
RP STRAIN=K12 / MG1655;
RC MEDLINE=97426617; PubMed=9278503;
RX Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=06:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesech P.,
RA Rascko D., Buckles E.L., Lion S.-R., Boutin A., Hackett D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli."
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
CC -1- FUNCTION: UPTAKE OF GALACTOSE ACROSS THE BOUNDARY MEMBRANE WITH
-1- THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
DR EMBL, U28377, AAA69110.1, -;
DR EMBL, AE000377, AAC75980.1, -;
DR EMBL, AE016766, AAN81977.1, ALT INIT.
DR PIR, F65079, F65079.
DR Ecogene; EGI2148; galP.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003663; Sug_transport.
DR Pfam; PF00083; sugat tr; 1.
DR PRINTS; PR00171; SUGRTNSPORT.
DR TIGRfam; TIGR00879; SP; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport;
KW Complete proteome.
FT DOMAIN 1 15 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 16 36 1 (POTENTIAL).
FT DOMAIN 37 56 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 57 77 2 (POTENTIAL).
FT DOMAIN 78 84 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 85 105 3 (POTENTIAL).
FT DOMAIN 106 112 PERIPLASMIC (POTENTIAL).

FT TRANSMEM 113 133 4 (POTENTIAL).
FT DOMAIN 134 139 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 140 160 5 (POTENTIAL).
FT DOMAIN 161 171 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 172 192 6 (POTENTIAL).
FT DOMAIN 193 250 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 251 271 7 (POTENTIAL).
FT DOMAIN 272 290 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 291 311 8 (POTENTIAL).
FT DOMAIN 312 321 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 322 342 9 (POTENTIAL).
FT DOMAIN 343 351 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 352 372 10 (POTENTIAL).
FT DOMAIN 373 394 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 395 415 11 (POTENTIAL).
FT DOMAIN 416 416 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 417 437 12 (POTENTIAL).
FT DOMAIN 438 464 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 464 AA; 50982 MW; 07E08935BD8E3F8B CRC64;
Query Match 11.1%; Score 451; DB 1; Length 464;
Best local similarity 21.6%; Pred. No. 5.5e-22;
Matches 160; Conservative 88; Mismatches 183; Indels 310; Gaps 16;
QY 62 VAIVASIGNILQGMNDNTIAAVALYIKKEFQLONEPTVEGLIVGMSLIGATVITTFSGPL 121
DB 18 VCFIAAAGLLFGIDIGVIALGLPFIADRFITSH--TQENVVSSMGAAVAGVSGML 75
QY 122 SDSIGRRPMLISSILYFFSGILMELMSBNVYLLARFVDFGIGLAVTLVLYSEIAP 181
DB 76 SFKIGRRKSLMIGAILFVAGSLFSAAPNVVLLISRLVLGLAVGVSATPLVYSEIAP 135
QY 182 SEIGGLNTLPQFSGSGCMFLSYCMVRCMSLSPSPDMYMGVLAIPSLFFGLITFLP 241
DB 136 EKIRGSMISMYOLMITTIGILGAY--LSDTASVYGAMWMLGVITIPILIL--IGVFLLP 192
QY 242 ESPRNLVSKGMAABAKVYLQLRGKDVSGELSLLEGLVGCPTSIEEYIIGPATEAD 301
DB 193 DSPRNFAPAKRFVDAERVLRLR-----DTSAP----- 220
QY 302 DLVTDDKREQITLYGPEEGSWIAPSKPIMLGSLVSLASRHSWMVNSVPLMDPIVTL 361
DB 221 ----- 220
QY 362 FGSVHNNPQAGGSMRSTLPFNFGSMFVTDHAKNEQWDEENLHRDDEYASDAGSDY 421
DB 221 -----AKREL----- 225
QY 422 EDNLHSPILSRQATGAEKQDVIHHGRGSLSMRQTLLEGSGDVSSTDIGGQQLAMK 481
DB 226 -DEIRSLQVKOS----- 241
QY 482 WSEKENGRRKEGKRVYLHQEGVPGSRGSIYSLPGGQDVFRGSEFVHAALVSGAL 541
DB 242 -----FK-----ENSNP----- 248
QY 542 FSKGLAEPKMDAMVHSEVAAGSRWKDLPBGVRRALLVGVGIQLQOPAGINSVLY 601
DB 249 -----RAVFLGVLLQVMOQFTGNVITV 272
QY 602 YTPQILQAGVAVILSKFGLSASASILISSLTLLMPCGFMALMDLSGRFLILGT 661
DB 273 YAPKIFELAGTNTTEQWM-----GTIVGLTNVLTATIAIG--LVDRGRK----- 317
QY 662 IPIILASVILVSVNLIDIGTLAHLSTVS---VIVYFCFPWGF---GPIPNILCA 713
DB 318 -PTLTLGLVMAAGWV-LGTMHGHISPSAQFALMLMFIYGFAMSGPLITVWCS 375
QY 714 EIPFRVAGLCIALCAFPWIGDIIYVTSLEPVMNLALGLAVFSIYAVVCLISFVFLK 773
DB 376 EIOPKGDGFIQTSSTANWLANMVGATFLMTLNTIGNANTFWYALANVLFILLTLM 435
QY 774 VPETKGMPLVETTFEFAVGAK 794

DB	436	VPEKGVSLSEHIERMLMKGRK	456
RESULT 7			
AAAE_KLEOX	STANDARD;	PRT;	472 AA.
AC P45558;			
DT 01-NOV-1995	(Rel. 32, Created)		
DT 01-NOV-1995	(Rel. 32, Last sequence update)		
DT 28-FEB-2003	(Rel. 41, Last annotation update)		
DE Arabinose-proton symporter (Arabinose transporter).			
GN AAAE.			
OS Klebsiella oxytoca.			
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
OC Enterobacteriaceae; Klebsiella.			
OX NCBI_Taxid=571;			
RN (1)			
RP SEQUENCE FROM N.A.			
RC STRAIN=8017;			
RX MEDLINE=9539486; PubMed=7665532;			
RA Shatwell K.P., Charalambos B.M., McDonald T.P., Henderson P.J.F.;			
RT "Cloning, sequencing, and expression of the <i>araB</i> gene of <i>Klebsiella</i>			
RT <i>oxytoca</i> 8017, which encodes arabinose-H ⁺ symport activity.";			
RL J. Bacteriol. 177:5379-5380(1995).			
CC -1- FUNCTION: UPAKE OF ARABINOSE ACROSS THE BOUNDARY MEMBRANE WITH			
CC THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).			
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.			
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.			
CC -----			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC the European Bioinformatics Institute. There are no restrictions on its use by			
CC non-profit institutions as long as its content is in no way			
CC modified and this statement is not removed. Usage by and for commercial			
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC or send an email to license@isb-sib.ch).			
CC -----			
DR EMBL; X79598; CAAS6110.1; -.			
DR PIR; S47089; S47089.			
DR InterPro; IPR007114; MFS.			
DR InterPro; IPR005828; Sub_transporter.			
DR InterPro; IPR005829; Sug_transporter.			
DR InterPro; IPR003663; Sug_transporter.			
DR Pfam; PF00083; sugar_tr; 1.			
DR PRINTS; PR00171; SUGETRANSPORT.			
DR TIGRfams; TIGR00879; SP. 1.			
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.			
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.			
FW TranspOrt; Sugar transport; Transmembrane; Inner membrane; Symport.			
KW DOMAIN 1 29			
FT TRANSMEM 30 50			
FT DOMAIN 51 63			
FT TRANSMEM 64 84			
FT DOMAIN 85 91			
FT TRANSMEM 92 112			
FT DOMAIN 113 114			
FT TRANSMEM 115 135			
FT DOMAIN 136 154			
FT TRANSMEM 155 175			
FT DOMAIN 176 178			
FT TRANSMEM 179 199			
FT DOMAIN 200 257			
FT TRANSMEM 258 278			
FT DOMAIN 279 297			
FT TRANSMEM 298 318			
FT DOMAIN 319 325			
FT TRANSMEM 326 346			
FT DOMAIN 347 361			
FT TRANSMEM 362 382			
FT DOMAIN 383 404			
FT TRANSMEM 405 425			
FT DOMAIN 426 427			
FT PERIPLASMIC (POTENTIAL).			

FT	TRANSMEM	428	446	12 (POTENTIAL).
FT	DOMAIN	447	472	CYTOLASMIC (POTENTIAL).
SEQ	SEQUENCE	472 AA;	51732 MW;	410021E1BEE3D96E CRC64;
	Query Match	10.7%;	Score 434.5;	DB 1; Length 472;
	Beet Local Similarity	21.4%;	Pred. No. 6.6e-21;	
	Matches .158;	Conservative	98;	Mismatches 182; Indels 299; Gaps 14
QY	62	VAIVAISGNLLIQGMNATIAAAVLYIKKEFQLONEPTYEGGLIVKSLIGATITVTFSGPL	121	
DB	25	VSIAAANVAGLLFGDIDIGVAGALPFIIDHFLYLSR--LQEWVSSMMLGAALGFNGL	82	
QY	122	SDSICRRMLLISILYFPGSGILIMMSDNYVLLARFVDEGIGLAATLVPLYSIETAP	181	
DB	83	SFRIGRKSLIHWGAVLFAVGVSGSAFAISVENMLVARVLYGVAAGIASATYAPLYSEMS	142	
QY	182	SEIRGLNTLTPQFSQSGMPLSYCVAFVPMSTISPPSDMRIMLVLAIPSLFFPGTLTFYLP	241	
DB	143	ENVGSKMISMVQLMWITLGIWAAF--LSDTAISYSGNMWAMGLVALPAVLLIIIVIF-LP	199	
QY	242	ESPRMLVSKGMAEAKVYLQLRGKDVSGBELSLLEGLVWGDSITSEYIIIGPATEAD	301	
DB	200	NSPRLAEKGRHVAEEVLRMLR-----DTS-----EKARD	230	
QY	302	DLVYDGDGEQITLYIGPEEGQGMIRPSPKPIMLGSVLSLARHSGMNVQSPVLMPIYTL	361	
DB	231	EL-----NEISESLKQ-----	243	
QY	362	FGSVHENPQAGSWSRTLFPNFGSMFVSTDQHAKNQWDEENLHRDDEEYASDAGGDY	421	
DB	244	-----	243	
QY	422	EDNLHSPILISQATGAEGKDIVHHGRGSALSMRRQTLLEGSGDGVSSSTDIQGGWQLAMK	481	
DB	244	-----GGW-----	246	
QY	482	MSEKENGKREGGKFRVYLLHQEGVPSGRBSIYSLPGGQGVFEGSEFVHAALVQSAL	541	
DB	247	-----AL	248	
QY	542	FSKGLAEPBMSDAAWHPSEVAAGSRWKDLFEPGVRALLVGVCIILQOPAGINGVLY	601	
DB	249	F-----KYNR-----NRRRAVFLGMLLQMQQFTGNITMY	279	
QY	602	YTPQILEQAGAVILSKFGLSSASASIISSLTLLMLPCIGFAMLLMDLSGR-FLLLG	660	
DB	280	YAPRIFFKMA-----GFTTQGGQVAVTLVGLTFMEAFVIAVFTVDKXGRKALKIG	330	
QY	661	TIPILIASVLIVVSNLIDGLTAHALISTVSVIYFCC--FWMGEPPIINILCAELFP	717	
DB	331	FSWAIAIGTLVIGYCLMQFDNGT-ASSGISMVSGVTMNCIAGYAMSAAPVWIIICSEIOP	389	
QY	718	TRVRLGLCAICAFTEWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFLKVPET	777	
DB	390	LKCDPFGITCTTTNWSNMIIIGATFTLLDIALGAGTFMLITYALTALNVAFIGVTFMLIPET	449	
QY	778	KGMPLEVIITEFPAVGAK	794	
DB	450	KNVTLLEHIERLRMSGK	466	
RESULT 8				
XYLE_ECOLI				
AC	P03058;	STANDARD;	PRT;	491 AA.
DT	01-MAR-1989 (Rel. 10, Created)			
DT	01-MAR-1989 (Rel. 10, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	D-xylose-proton symporter (D-xylose transporter).			
GN	XYLE OR B4031 OR Z5629 OR EC55014.			
OS	Escherichia coli, and			
OS	Escherichia coli O157:H7.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			

CC Enterobacteriaceae; Escherichia.
CX NCBI_TaxID=562, 83334;
RN RP
RN SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=88007632; PubMed=2820984;
RA Davis E.O., Henderson P.J.F.;
RT "The cloning and DNA sequence of the gene xylE for xylose-proton
RT symport in Escherichia coli K12.";
RL J. Biol. Chem. 262:13928-13932(1987).
RN RN
RN SEQUENCE FROM N.A.
RP MEDLINE=87115869; PubMed=3543693;
RX Maiden M.C.J., Davis E.O., Baldwin S.A., Moore D.C.M.,
RA "Henderson P.J.F.;
RT "Mammalian and bacterial sugar transport proteins are homologous.";
RL Nature 325:641-643(1987).
RN RN
RN SEQUENCE FROM N.A.
RP STRAIN=K12 / MG1655;
RX MEDLINE=94089392; PubMed=8265357;
RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
RA Daniels D.L.;
RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the
RT region from 89.2 to 92.8 minutes.";
RL Nucleic Acids Res. 21:5408-5417(1993).
RN RN
RN SEQUENCE FROM N.A.
RP STRAIN=O157:H7 / EDL933 / ATCC 700927;
RC MEDLINE=21074935; PubMed=11206551;
RX Perna N.T., Plunkett G. III, Burland V., Mau B., Glaser J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postif G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potaomousis K.,
RA Apodaca J., Anantharaman T.S., Ian J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7,"
RL Nature 409:529-533(2001).
RN RN
RN SEQUENCE FROM N.A.
RP STRAIN=O157:H7 / RIMD 0509952;
RC MEDLINE=21156231; PubMed=11258796;
RX Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsuda E., Nakayama K., Muraoka K., Tanaka M., Tobe T.,
RA Ikida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
RA Kishida S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN RN
RN SEQUENCE OF 1-192 FROM N.A.
RP STRAIN=K12;
RX MEDLINE=88234001; PubMed=2836810;
RA Francoz E., Dassa E.;
RT "3' end of the malFG operon in E.coli: localization of the
RT transcription termination site.";
RL Nucleic Acids Res. 16:4097-4109(1988).
CC CC
CC -1- FUNCTION: UPAKE OF D-XYLOSE ACROSS THE BOUNDARY MEMBRANE WITH
CC THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -1- INDUCTION: By xylose.
CC -1- MISCELLANEOUS: E. COLI HAS TWO D-XYLOSE TRANSPORT SYSTEMS THAT
CC ACCUMULATE SUGAR AGAINST A CONCENTRATION GRADIENT: THE XYLE SYSTEM
CC WHICH UTILIZES THE ELECTROCHEMICAL GRADIENT OF PROTONS AND THAT IS
CC INSENSITIVE TO COLD OSMOTIC SHOCK AND THE XYLE SYSTEM THAT USES A
CC HIGH-ENERGY PHOSPHATE COMPOUND AND IS SENSITIVE TO COLD OSMOTIC
CC SHOCK.
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; J02812; AAA79016.1; -
DR	EMBL; U00006; AAC43125.1; -
DR	EMBL; AE000476; AAC77001.1; -
DR	EMBL; AE005636; AAG59230.1; -
DR	EMBL; AP002568; BAB38437.1; -
DR	EMBL; X06663; CAA29863.1; -
DR	PIR; A26430; A26430.
DR	PIR; B86096; B86096.
DR	PIR; F91255; F91255.
DR	Ecocore; EG1076; xyle.
DR	InterPro; IPR007114; MES.
DR	InterPro; IPR005828; Sub transporter.
DR	InterPro; IPR005829; Sug transporter.
DR	InterPro; IPR003663; Sugar transp.
DR	Pfam; PF00083; sugar tr; 1.
DR	PRINTS; PR00171; SUGRTNSPOT.
DR	TIGRPFAMs; TIGR00879; SP; 1.
DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW	transport; Sugar transport; Transmembrane; Inner membrane; Symport;
KW	Complete proteome.
FT	DOMAIN 1 9 CYTOPLASMIC (POTENTIAL).
FT	TRANSSEM 10 30 1 (POTENTIAL).
FT	DOMAIN 31 55 PERIPLASMIC (POTENTIAL).
FT	TRANSSEM 56 76 2 (POTENTIAL).
FT	DOMAIN 77 89 CYTOPLASMIC (POTENTIAL).
FT	TRANSSEM 90 110 3 (POTENTIAL).
FT	DOMAIN 111 133 PERIPLASMIC (POTENTIAL).
FT	TRANSSEM 134 154 4 (POTENTIAL).
FT	DOMAIN 155 165 CYTOPLASMIC (POTENTIAL).
FT	TRANSSEM 166 186 5 (POTENTIAL).
FT	DOMAIN 187 200 PERIPLASMIC (POTENTIAL).
FT	TRANSSEM 201 221 6 (POTENTIAL).
FT	DOMAIN 222 272 CYTOPLASMIC (POTENTIAL).
FT	TRANSSEM 273 293 7 (POTENTIAL).
FT	DOMAIN 294 312 PERIPLASMIC (POTENTIAL).
FT	TRANSSEM 313 333 8 (POTENTIAL).
FT	DOMAIN 334 343 CYTOPLASMIC (POTENTIAL).
FT	TRANSSEM 344 364 9 (POTENTIAL).
FT	DOMAIN 365 369 PERIPLASMIC (POTENTIAL).
FT	TRANSSEM 370 390 10 (POTENTIAL).
FT	DOMAIN 391 407 CYTOPLASMIC (POTENTIAL).
FT	TRANSSEM 408 428 11 (POTENTIAL).
FT	DOMAIN 429 442 PERIPLASMIC (POTENTIAL).
FT	TRANSSEM 443 463 12 (POTENTIAL).
FT	DOMAIN 464 491 CYTOPLASMIC (POTENTIAL).
FT	CONFUCT 64 64 A-> V (IN REF. 6).
FT	SEQUENCE 491 AA; 53608 MW; 2AF1AP9756C0B722 CRC64;
QY	Query Match 10.7%; Score 434; DB 1; Length 491;
QY	Best Local Similarity 19.5%; Pred. No. 7.5e-21;
QY	Matches 152; Conservative 109; Mismatches 175; Indels 344; Gaps 17
QY	62 VAIVASIGNLILQGMNDNATIAAATLYIKKEF---QLQNEPTVEGLI---VNSSLIGATVTT 115
QY	13 ITLVATITGLLFGFDYAVISGTVESLNTVPFAAPQMLSESANSLGFCVASALIGCIIGG 72
QY	116 TFSGGLPSIGIRRRPMLTILSSILVEFSGIIMLW-----SPN-----YVV-----LLLA 157
QY	73 ALGGYCSRRFRKRSKLIAALVFLTISGGSAMPGLGFSINPDNTVPYILAGYVEFVIY 132
QY	158 RFVDGFGIGLAVTTLVPLIYISIEASEIRGLANTLTPQFSGSGMPLSYCMVFGMSIS---- 213
QY	133 RIIGIGIGVGLASMLSPMYIAELAPAHIRKGLVSNFGAIFIGQLLVYCVNYFIARSGDAS 192
QY	214 --PSPDWRIMLGVLAIPSLFFPGGLITPLPSSPRMLVSKGMAEKAKYLOKLRGDDYVG 271
QY	193 WLANDGWMYMASBCIPALLFL-MLLYVPPSSPRMLVSKGQEQAGILRKIMGT----- 247

FT	TRANMEM	469	489	11 (POTENTIAL).
FT	DOMAIN	498		EXTRACELLULAR (POTENTIAL).
FT	TRANSEM	499	519	12 (POTENTIAL).
FT	DOMAIN	520	557	CYTOPLASMIC (POTENTIAL).
FT	VARIANT	271	271	N -> H (IN STRAIN PR745).
FT	VARIANT	428	428	N -> H (IN STRAIN PR745).
FT	VARIANT	436	436	L -> I (IN STRAIN PR745).
FT	VARIANT	526	526	S -> F (IN STRAIN PR745).
EQ	SEQUENCE	557 AA;	6136 MM;	20875EC1B153175 CRC64.

Query Match	Similarity	10.6%	Score 432	DB 1	Length 557
Best Local Similarity	20.5%	Pred. No. 1.2e-20			
Matches	166	Conservative 114	Mismatches 212	Indels 318	Gaps 22
Qy	10	QTPPTDLDRRERLLPSVVALPGLPPASCSGQEPVTSDDLE---	DKMSGAV-LVAI	64	
Db	35	KVPFSEDM---REIPI---LPNEEANAATPDQANEVADENGEGFEAKTISMIWVISA	184	86	
Qy	65	VASIGNLLQGMNATIAAAVLYIKKEFQLONEPTVEGLIVSNLIGATITVTTSGPLSDS	124		
Db	87	VAGISGLLFEGYDTGVISGALVAGSDLGHVILSSGCKELTATSPPAALISATTSQMLDW	146		
Qy	125	IGRRPMLLSSTLYFSGILMMSNNVYLLARFVDDGIGLTAUPLVPLYSILAPSEI	184		
Db	147	VGRKRLLCADAIPIYIGSVIIMASNNVAMVNGRIYVIGILTSILVPMYITELAPAL	206		
Qy	185	RGLNTLTPQFSGSGGMPFLSYCMVFGMSLPS-----PWRIMLVIAIPLSEFPFG--LTI	237		
Db	207	RGRVLIIVVFITGQLIAY-----SLMAAEHVHOGMRIMFGIAPAL---GQLSL	257		
Qy	238	FYLPSPFMYLVSKGMAEAKKYLQLRKGDVDVSGELSLLSGLVEVGDTISIEYITIGPAT	297		
Db	258	FMTPESPRYLLRHNNHVEKYKYLISIH-----	284		
Qy	298	EAADLDVTDGKQKITLYGPEEGGQMIARPSKPGIMLSGVSLSARHSGMNVGSLPMDP	357		
Db	285	-----PE-----AKPA-----	290		
Qy	358	IVTLFGSVHENMPQAGSNRSTLPNFGSMFSTVDQHAKNQOMDEENLRHDEEYASDGA	417		
Db	291	-----	290		
Qy	418	GGDYEDNLHSPILSRQATGABGKDIVHHGHSALSMRRQTLIGEGGDGVSSTDIGGMQ	477		
Db	291	-----	291		
Qy	478	LAMKSEKEGNGRKEGGFKRYVLLHQEGVPSGRBSIYSLPGGQGVFEGSEFVHAALVS	537		
Db	292	IAYK-----VSLIQEGVK-----VDFP-----EENKROH-----	315		
Qy	538	QSALFESKGLAEPBMSDAAWVHPSEVAAKGRKDKLFE--PGVRRALLVGGIILQOFAI	596		
Db	316	-----PFHPL-----KULFVPSNRSRSLFICGFLQWPOQFSST	348		
Qy	597	NGVLVYTPQILEQAGVAVILSKFGLSNASASILISLTTLLMPCIGFAMLMIDLSGRRF	656		
Db	349	NAIQFSAIIFQSVGFK-----NSISVSIVGATNVPFTI---VAFMFIDRIGRRR	396		
Qy	657	LLLGITPILIASVLIVVS-NLIDGLTAHA-----LLSTVAVIYFCGFVNGFGPIPI	710		
Db	397	ILLCSAVMIAGLACAIAYHFLPADTTONTSNGMOYVVLASITILFLASVASIGINIP-W	455		
Qy	711	LCABIFPTPRVRLCIAICAFTEWIDDIYTVSLPVMNAIGLAGVFSIYAAVCLISFVNV	770		
Db	456	QQAELFPMHEVRLAGGFSFALNMVGNLLISAFFLTMBSITPTGTFFALPAGCFPGVLVS	515		
Qy	771	FLKVPETKGMPLVEVITEFPFVAGAKQAAAKA	800		
Db	516	YFTYPELAGMSIENIHKLLEKFGWQAVKES	545		

```

ID  GTRI.LEIDO      STANDARD;      PRT;      547 AA.
AC  Q01440;
DT  01-JUN-1994 (Rel. 29, Created)
DT  01-JUN-1994 (Rel. 29, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Membrane transporter D1.
OS  Leishmania donovani.
OC  Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX  NCBI_TaxID=5661;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=93063058; PubMed=1435876;
RA  Langford C.K., Embank S.A., Hanson S.S., Ullman B., Landfear S.M.;
RT  "Molecular characterization of two genes encoding members of the
RT  glucose transporter superfamily in the parasitic protozoan Leishmania
RT  donovani.";
RL  Mol. Biochem. Parasitol. 55:51-64(1992).
CC  -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC  -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M85072; AAA29230.1; -.
DR  PIR; A48442; A48442.
DR  InterPro; IPR007114; MFS.
DR  InterPro; IPR005828; Sub_transporter.
DR  InterPro; IPR005829; Sug_transporter.
DR  InterPro; IPR003663; Sugar_transpc.
DR  Pfam; PF00083; sugar_tr; 1.
DR  PRINTS; PRO0171; SUGRTNSPORT.
DR  TIGRFAWS; TIGR00879; SP; 1.
DR  PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR  PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW  Transport; Sugar transport; Transmembrane.
FT  DOMAIN 1
FT  TRANSMEM 3 25
FT  DOMAIN 26 43
FT  TRANSMEM 44 64
FT  DOMAIN 65 78
FT  TRANSMEM 79 99
FT  DOMAIN 100 101
FT  TRANSMEM 102 122
FT  DOMAIN 123 136
FT  TRANSMEM 137 157
FT  DOMAIN 158 164
FT  TRANSMEM 165 185
FT  DOMAIN 186 245
FT  TRANSMEM 246 266
FT  DOMAIN 267 276
FT  TRANSMEM 277 297
FT  DOMAIN 298 308
FT  TRANSMEM 309 329
FT  DOMAIN 330 339
FT  TRANSMEM 340 360
FT  DOMAIN 361 385
FT  TRANSMEM 386 406
FT  DOMAIN 407 428
FT  TRANSMEM 429 547
FT  DOMAIN 547 547
SQ  SEQUENCE 547 AA; 58787 MW; CA8EEF16EB0FF514 CRC64;

Query Match 10.6%; Score 431; DB 1; Length 547;
Best Local Similarity 19.7%; Pred. No. 1.3e-20;
Matches 146; Conservative 101; Mismatches 178; Indels 316; Gaps 12;

```

```

Db      5 VMLCALGGLFGYDVTGYNALPQMKDFGFSHEHWOYALVIAIAGAFVAGFISGFI 64
Qy      122 SDISGRPMMLISLTFFSGILMSPVYLLARFPVDFGIGLAVTVLTVLSEIAP 181
Db      65 SAARFRRRCIAADALFVIGSVLMAAPVVEVVLVSRVVGAIIGISATTPVYLAETVS 124
Qy      182 SEIRGLNLTLPQFSGGWFLS--YCNVFGMSLSPDWRIMLVYLAIPSLFFGLTTFY 239
Db      125 PKHRGATIVLNNLFLTGGQFVAAGFTALNVFTSNIGMRVAILGALPAVVAQACLLPFF 184
Qy      240 LPESRRLVSKGRMAEAKVYLOKRGKDDVSGELSLLEGLVEGDDTSIEYITGPATBA 299
Db      185 LPESRRLVSKGHADRAVAV----- 204
Qy      300 ADDLTDDGKEQITLYGPEEGSWIARPSKPIMLJSVLSLASHRGSMVNSVPLMDPIV 359
Db      205 ---ADKTEVDLCFEPEG----- 218
Qy      360 TLFGSVHNMPOAGSGMSRSTLFPNFGSMFSVTDQAKNQWDEENLHDEEYASDGAGG 419
Db      219 ---DELPL---SVRL----- 226
Qy      420 DYEDNLHSPPLSRQATGABKDIYHHGHRGALSMRQOTLLGEGGDVSGSTDIGGKQOLA 479
Db      227 DYR---PLMARL----- 235
Qy      480 WKMSKEBENRGKGGFKRVYVHOGVPGSRGSIYSLPGGDVFESEFVHAALVVSQS 539
Db      236 --- 235
Qy      540 ALFSGKLAEPMSDAMVHPSEVAAKGRKDLFEPGVRAVLVGVGIQLQCPAGINGV 599
Db      236 ---NRFRVLSGGIQLQCPAGINTI 258
Qy      600 LYTTPQILBQGVAVILSKFGSSASLISLTLLMLPCTIGRAMLMDISGRFILL 659
Db      259 MYSSVIIYDAGFRDAIMPVVISIPLA--FMNALFTAV---AIFVDRFGRRRML 309
Qy      660 GTIPPLIASVLYVNSLIDL-----GTLAHLSTVSIVVFCPCVMGFRP 708
Db      310 ISV---PGLVLLVVAIIIGFTGRISYVGGGLFLALLA---VFALAPDIGICP 361
Qy      709 NILCAIEPFRVRLCIAICAFTEWIGDIIVYSLPVMNLAIAGLGVSIYAVCLISFV 768
Db      362 WVMGEIPEPHTLRISAASVATMANNGAVLVSGVPIIMGALGVGFTIISGLMALGCI 421
Qy      769 FVFLKVPETKGMPLVEITTEFF 789
Db      422 FVYFPVAVETKGLTEQIDNMF 442

RESULT 11
AAAE_ECOLI STANDARD; PRT; 472 AA.
ID      AAEE_ECOLI
AC      P09830; Q46937;
DT      01-MAR-1989 (Rel. 10, Created)
DT      01-MAR-1989 (Rel. 10, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Arabinose-proton symporter (Arabinose transporter).
GN      ARAE OR B2841 OR Z4161 OR ECS3698.
OS      Escherichia coli, and
OS      Escherichia coli O157:H7.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Escherichia.
OX      NCBI_TaxID=562, 83334;
RX      MEDLINE=87115869; PubMed=3543693;
RA      Maiden M.C.J., Davis E.O., Baldwin S.A., Moore D.C.M.,
RA      Henderson P.J.F.,
RA      "Mammalian and bacterial sugar transport proteins are homologous.",
RL      Nature 325:641-643 (1987).
RN      [2]

```

```

RP      SEQUENCE FROM N.A.
RC      STRAIN=K12 / JM2433;
RX      MEDLINE=88228015; PubMed=2836407;
RA      Maiden M.C.J., Jones-Mortimer M.C., Henderson P.J.F.;
RT      "The cloning, DNA sequence, and overexpression of the gene araE
RL      coding for arabinose-proton symport in Escherichia coli K12.";
RN      J. Biol. Chem. 263:8003-8010 (1988).
RP      SEQUENCE FROM N.A.
RC      STRAIN=K12 / MG1655;
RX      MEDLINE=97426617; PubMed=9278503;
RA      Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA      Riley M., Collado-VIDE J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA      Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA      Mau B., Shao Y.;
RT      "The complete genome sequence of Escherichia coli K-12.";
RN      Science 277:1453-1474 (1997).
RP      SEQUENCE FROM N.A.
RC      STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX      MEDLINE=21074935; PubMed=11206551;
RA      Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA      Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA      Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA      Grobeck B.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA      Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA      Welch R.A., Blatter F.R.;
RT      "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RN      Nature 409:529-533 (2001).
RP      SEQUENCE FROM N.A.
RC      STRAIN=O157:H7 / RIMD 0509952;
RX      MEDLINE=21156231; PubMed=11259796;
RA      Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA      Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA      Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
RA      Kihara S., Shiba T., Hattori M., Shinagawa H.;
RT      "Complete genome sequence of enterohaemorrhagic Escherichia coli
RL      O157:H7 and genomic comparison with a laboratory strain K-12.";
RN      DNA Res. 8:11-22 (2001).
RP      PRELIMINARY SEQUENCE OF 1-28 FROM N.A.
RX      MEDLINE=84114868; PubMed=6319708;
RA      Stoner C., Schleif R.F.;
RT      "The araE low affinity L-arabinose transport promoter. Cloning,
RT      sequence, transcription start site and DNA binding sites of
RL      regulatory proteins.";
RN      J. Mol. Biol. 171:369-381 (1983).
RP      -I- FUNCTION: UPTAKE OF ARABINOSE ACROSS THE BOUNDARY MEMBRANE WITH
CC      THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).
CC      -I- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC      -I- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; J03732; AAA23469.1; -
DR      EMBL; X00272; CAA25075.1; -
DR      EMBL; U29581; AAB40488.1; -
DR      EMBL; AE000368; AAC75880.1; -
DR      EMBL; AE005113; AAC57953.1; -
DR      EMBL; AP002563; BAB37121.1; -
DR      PIR; B26430; B26430.
DR      PIR; B91091; B91091.
DR      PIR; B85936; B85936.
DR      EcoGene; EG10056; araE.
DR      InterPro; IPR007114; MFS.
DR      InterPro; IPR005828; Sub_transporter.

```

DR InterPro; IPR005829; Sug_transporter.
 DR InterPro; IPR003663; Sugar_transp.
 DR Pfam; PF00083; sugar_tr; 1.
 DR PRINTS; PR00171; SUGTRNSPRT.
 DR TIGRFAMs; TIGR00679; SP; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport;
 Complete proteome.
 FT DOMAIN 1 29 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 30 50 1 (POTENTIAL).
 FT DOMAIN 51 63 PERIPLASMIC (POTENTIAL).
 FT TRANSMEM 64 84 2 (POTENTIAL).
 FT DOMAIN 85 91 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 92 112 3 (POTENTIAL).
 FT DOMAIN 113 114 PERIPLASMIC (POTENTIAL).
 FT TRANSMEM 115 135 4 (POTENTIAL).
 FT DOMAIN 136 154 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 155 175 5 (POTENTIAL).
 FT DOMAIN 176 178 PERIPLASMIC (POTENTIAL).
 FT TRANSMEM 179 199 6 (POTENTIAL).
 FT DOMAIN 200 257 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 258 278 7 (POTENTIAL).
 FT DOMAIN 279 297 PERIPLASMIC (POTENTIAL).
 FT TRANSMEM 298 318 8 (POTENTIAL).
 FT DOMAIN 319 325 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 326 346 9 (POTENTIAL).
 FT DOMAIN 347 361 PERIPLASMIC (POTENTIAL).
 FT TRANSMEM 362 382 10 (POTENTIAL).
 FT DOMAIN 383 404 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 405 425 11 (POTENTIAL).
 FT DOMAIN 426 427 PERIPLASMIC (POTENTIAL).
 FT TRANSMEM 428 446 12 (POTENTIAL).
 FT DOMAIN 447 472 CYTOPLASMIC (POTENTIAL).
 FT CONFLICT 26 28 SVA -> YDR (IN REF. 6).
 SQ SEQUENCE 472 AA; 51684 MW; 411904441D44393 CRC64;
 Query Match 10.5%; Score 429.5; DB 1; Length 472;
 Best Local Similarity 20.6%; Pred. 1.4e-20;
 Matches 154; Conservative 100; Mismatches 194; Indels 299; Gaps 13;
 QY 52 LEDKMSGAVVAIVASIGNLLQGMNDATIAAIVYIKKEFOLQNEPTEGLIVSNLIGA 111
 DB 15 LQDTRRMKMFVVAAGVGLFGLDVGIVAGLPTTHFVLTSL-QQEWVSSMMLGA 72
 QY 112 TIVTTFSPPLSDISGRPMILISLILYFPGSLIMSPNVYLLARFYDGFGLAVTL 171
 DB 73 AIGALFNGMLSPRLGRKYSILMAGALIFVIGSIGSAFATSVEMLIARVVLGIIVGIAST 132
 QY 172 VPIVSEIAPSEIRGLMTLTPQFSGSGMFLSYCVFPGMSLSPSPDWRIMIGVLAIPSLF 231
 DB 133 APLVISEASENVRKMTSMQIMVTLGIVLAF-LSDTAYSYSGNMRAMGLVLLPAVL 190
 QY 232 FFGTLTIPYLPSPFWLVSGRMAEAKVYLQKLRGDVSGELSLLEGLVGGDTSEY 291
 DB 191 LILLVLF-LPMSPRFLAEKRHIEAEVILPMR-----DTS----- 225
 QY 292 IIGPATADLVLDGDKEQITLTYPEREGQSWIARPSKPIMLGSLVSLASHRSGMVNQS 351
 DB 226 -----EKARBEI-----NEIRIS 238
 QY 352 VPLMDPIYTLFGSVHNNPQAGSMRSTLFPNFGSMFSTVDQAHANQEDENLRDDE 411
 DB 239 LKLKO----- 243
 QY 412 YASDAGGDYEDNLHSLPLSRQATGAEKQIVHGHGSLSMRRQTLLEGEGDVSSSTD 471
 DB 244 ----- 243
 QY 472 IGGGQQLAMKSEKGEKNGRKGGRFRRVYLHOBGVGSRGSIIVLPGGGDVFESEFVH 531
 DB 244 --GGW----- 246

QY 532 AALVQSALFSSKLAEPBMSDAAVHPSEVAAKSRWKDLFEBCVRRALLVGVGIOLQ 591
 DB 247 -----ALFK-----INNVRRAVFLGMLQAMQ 269
 QY 592 QPAGINGVLYTTPQILEQAGVAVLSKFGSSASALISLTLTLLMPCIGFAMLLMDL 651
 DB 270 QFTGMNIIIMYAPRIFPKA-----GFTTQCOMATLVGLTFEAFPIAVFTVDK 320
 QY 652 SGR-FLITGPIPLISIVLIVNSLIDGLTAAHLLSTVSVYVPC---FWNGRPI 707
 DB 321 AGRPALKIGFSSVALGLTVGYCLMOPDNT-ASSGLSWLSVQMTMCIAGVANSAPV 379
 QY 708 PNIICAEIFFPRVAGLCIACAFPMWIGDIIIVTSLPMTLNAIGLAVFSIYAVCLISF 767
 DB 380 VMIICSEIQLKCDPFGTCSTTNWVSNMIIIGATFLILDSIGAGTFWLYTALNAFV 439
 QY 768 VFVFLKVPETKMPLEVTTEFFAVGAK 794
 DB 440 GTFWLLPETGVNLTLEHIERKLMAGEK 466
 RESULT 12
 ITL1_YEAST
 ID ITL1_YEAST STANDARD; PRT; 584 AA.
 AC P30605;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Myo-inositol transporter 1.
 GN ITL1 OR YDR497C OR D9719.3.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OK NCBI_TaxID=4932;
 RN RN
 RX MEDLINE=91250431; PubMed=2040626;
 RA Nikawa J.-I., Tsukagoshi Y., Yamashita S.;
 RT "Isolation and characterization of two distinct myo-inositol
 RT transporter genes of Saccharomyces cerevisiae";
 RJ J. Biol. Chem. 266:11184-11191(1991).
 RL [2]
 RP SEQUENCE FROM N.A.
 RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E.,
 RA Berno A., Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
 RA Hunkeler-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D.,
 RA Mosedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,
 RA Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,
 RA Winant A., Yelton M., Botstein D., Davis R.W.;
 RA Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MAJOR TRANSPORTER FOR MYO-INOSITOL.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).

```

DR PRINTS; PR00171; SUGRTNSPORT.
DR TIGRFAMS; TIGR00879; SP: 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Sugar transport; Glycoprotein.
FT DOMAIN 1 81
FT TRANSMEM 182 102
FT DOMAIN 103 129
FT TRANSMEM 130 150
FT DOMAIN 151 163
FT TRANSMEM 164 184
FT DOMAIN 185 186
FT TRANSMEM 187 207
FT DOMAIN 208 215
FT TRANSMEM 216 236
FT DOMAIN 237 246
FT TRANSMEM 247 267
FT DOMAIN 268 349
FT TRANSMEM 350 370
FT DOMAIN 371 376
FT TRANSMEM 377 397
FT DOMAIN 398 400
FT TRANSMEM 401 421
FT DOMAIN 422 441
FT TRANSMEM 442 462
FT DOMAIN 463 486
FT TRANSMEM 487 507
FT DOMAIN 508 510
FT TRANSMEM 511 531
FT DOMAIN 532 584
FT CARBOHYD 371 44
FT COMPLECT 43
SQ SEQUENCE 584 AA; 63569 MW; 42543E30A102CC55 CR664;

Query Match 10.3%; Score 421.5; DB 1; Length 584;
Best Local Similarity 19.1%; Pred. No. 6.1e-20;
Matches 150; Conservative 121; Mismatches 195; Indels 321; Gaps 19;

QY 44 EPNVSD-----LIEDKMGAVL-VAVVASTGNTLLCGMDNTAAVLYIKKEQOLONE 96
DB 64 KPVNEDDTQSVWITNOSISPEITLTLTFVASTISGFMFGYDGYISSALISGTD--LPHK 121
QY 97 PTVEE---LIVMSLIGATIVTFEFGPLSDSISGRPMILISILYFESGLMLMSPNYV 153
DB 122 VLTVEKRIIVATSLGLALITISIFAGTADIRGRKCLMSLUMVIGAILLOVSAHTWQ 181
QY 154 LLARFVDGFGIGLAVTLVPLYSIASEIRGLNLTLPQFSGSGMPLSYCMVFGMSLS 213
DB 182 MAVGRLIMGFGVIGISLPLFISIAIPMIRGLTVINSILWLTGGQVAVGCGAGLVN- 240
QY 214 PSPDRIMLGVAISLFFFGITFYLPESPRLWVSKRMAAKVLOKRGKDVSGEL 273
DB 241 VNNGRILVGLSLPTAVQF--TCLCFPLDTPRYVMKDLAVATVLK--RSYDTSEI 297
QY 274 SLLEGLEVGSGDTSEIYIIGPATEAADLVTDGKEQITLYGPEGQSWIARPSKGPIM 333
DB 298 -----IERKV-----BELVT----- 307
QY 334 LGSVLSLASHRGSMVNSVPLMDPIVTLFGSVHENVPOAGSGWRSTLPPNFGMSFTDQ 393
DB 308 -----LNOSIP----- 313
QY 394 HAKNEQWDEBNLHRDEYASDAGGDYEDNINSLRSQATGAGKDIYHHGRGSL 453
DB 314 -----GKVV----- 317
QY 454 MRROTLLGEGDGVSTDIGGGMQLAMKSEKENGREGGFKVYLHOGVPGSRG 513
DB 318 ----- 317
QY 514 IVSLPGGDVFESEFVHAALVQSALFSKGLAEFRMSDAMVHPSEVAAGSRWKDLF 573
DB 318 -----PEKV-----MNTIK 326

```

```

QY 574 E-----PGVRALVGVGIQLQFAGINGVLYYTPQILEQGVAVILSKPGLSASATL 629
DB 327 ELHVPNSLRALITIGCGQALQQTGMNLSMIFSGITIFETVGR-----NSSAVSI 378
QY 630 ISSLTLLMLPCIGFAMLMNDLSGRFLLGTITILASVIL-----VSN 676
DB 379 VSGNPIFTL-----VAFPSIDIKRITLILGLGMPTALVCSIAFPLGIRKFDGAVAV 434
QY 677 LIDGTLAALLSVSYIVYFCFVWGFGPIPLCAEIFPTRRGICIAICATFTWGD 736
DB 435 VSSGFSFGWGVIIIVFIVFAFALGIVP--WQSELPQNRIGICTSYATATNAGS 493
QY 737 IIVYVSLVPMNATIGLGVFSIVAVCLISFVFLKVPETKMPLE-----VITEFPAVG 792
DB 494 LVIASTPLTMQNTTPAGTFAFPAGLSCLSTIFCPYCPFLSGLEBVOITLIDGRNIK 553
QY 793 AKQAAR 799
DB 554 ASKALAK 560

RESULT 13
MYCT HUMAN STANDARD; PRT; 629 AA.
ID MYCT HUMAN
AC 0960E2;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Proton myo-inositol co-transporter (Hmlt).
GN SLC2A13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RX MEDLINE=21391710; PubMed=11500374;
RA Ulfert M., Ibberson M., Horstberger J.-D., Chatton J.-Y.,
RT Riederer B.M., Thorens B.;
RT "Identification of a mammalian H(+)-myo-inositol symporter expressed
RT predominantly in the brain.";
RL EMO J. 20:4467-4477(2001).
CC -!- FUNCTION: H(+)-myo-inositol co-transporter. Can also transport
CC related stereoisomers.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Predominantly expressed in the brain.
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; AJ315644; CAC51116.1; -.
CC Genew; HGNC:15956; SLC2A13.
CC InterPro; IPR007114; MFS.
CC InterPro; IPR005828; Sug_transporter.
CC InterPro; IPR005829; Sug_transporter.
CC InterPro; IPR003663; Sugar_transpct.
CC Pfam; PF00083; sugar_tr; 1.
CC PRINTS; PR00171; SUGRTNSPORT.
CC TIGRFAMS; TIGR00879; SP: 1.
CC PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
CC PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Sugar transport; Glycoprotein.
FT DOMAIN 1 57
FT TRANSMEM 58 78
FT DOMAIN 79 106
FT TRANSMEM 107 126
SQ SEQUENCE 106 AA; 11569 MW; 42543E30A102CC55 CR664;

```

```

FT TRANSMEM 107 127 2 (POTENTIAL).
FT DOMAIN 128 129 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 130 150 3 (POTENTIAL).
FT DOMAIN 151 159 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 160 180 4 (POTENTIAL).
FT DOMAIN 181 193 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 194 214 5 (POTENTIAL).
FT DOMAIN 215 220 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 221 241 6 (POTENTIAL).
FT DOMAIN 242 305 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 306 326 7 (POTENTIAL).
FT DOMAIN 327 344 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 345 365 8 (POTENTIAL).
FT DOMAIN 366 374 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 375 395 9 (POTENTIAL).
FT DOMAIN 396 489 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 490 510 10 (POTENTIAL).
FT DOMAIN 511 530 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 531 551 11 (POTENTIAL).
FT DOMAIN 552 554 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 555 575 12 (POTENTIAL).
FT DOMAIN 576 629 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 414 414 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 439 439 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 629 AA; 68203 MW; 8AD79C56C790814E CRC64;

```

```

Query Match 10.3%; Score 421; DB 1; Length 629;
Best Local Similarity 18.5%; Pred. No. 7.2e-20;
Matches 151; Conservative 105; Mismatches 183; Indels 378; Gaps 14;

```

```

QY 61 LVATVATAGNLTGMDNMTAAVYIKKEPQLQNEPTVEGLIVMSLIGATVTTFS-G 119
DB 64 VVAVSALGGFPGIDTGVVSGAMLLK--OLSIDALMOELLVS-STYGAASALAGG 120
QY 120 PLSDSIGRRPMLLSIIYFSGSLIMLSPVNYVLLAFVDFGFGIGLAVTLVPLYSBI 179
DB 121 ALNGVFGRRATLLASALFTAGSAVLAANNKETILLAGLVGIGLIGASMTVPYIAEV 180
QY 180 APSEIRGLINTLPQSSGCGMFLSYCMFGMSLSPSPDRIMLVLAIPSLF-FGLTIF 238
DB 181 SPNNRGRGLVITINTLFTGCGPFASVVDCAFSLYKDKGRMYLGLAXYPAVYQFPG--FL 238
QY 239 YLPSPRMVSKGRMAKXVLQKLRKGDVSGELSLLEGLVEGSDISIEYIIIGPATE 298
DB 239 FLPSBPRLIQGQYOKARRILISQMG----- 265
QY 239 AADLVTDGDKQITLYGPEEGQSWIARPSKGPIMLGSVLSLASHSGMWSQVPLMDPI 358
DB 266 -----NOTI----- 269
QY 359 VTLFSGVHENMPQAGSMKSTLPFNFGMSFVTDHAKNEQND--EENIHRDEYASDG 416
DB 270 -----DEEYDSIKNNIEEKEVGSAG 291
QY 417 AGGDVEDNHSPLLSRQATGAEGKOIVHHHNGSALSMRQTLLEGGGVSTDIIGGM 476
DB 292 -----PVICM----- 297
QY 477 QLAWKSEKEGNGRKEGFKREVYLHQEGVPSRRGSIVSLPGGDVFGSEFVHAALV 536
DB 298 ----- 297
QY 537 SQSALFSKGLAEPMSDAAVHPSEVAAGSKMKOLFEBGVRRALLVGVGIILOQFAGI 596
DB 298 -----LSYPTRRALLVGGGLQMFQOLSGI 322
QY 597 NGVLYTTPQILBOAGVAVILSKFGLSSASASILISLTLLMLPCGFMMLMDLSGRPF 656
DB 323 NITMYSATILQMSGVED--DRILAWLASVTAFTNFTLV-----GWLVEKVRKX 373
QY 657 LLIGTPIILASLVILV----- 674

```

```

DB 374 LTFPSLACTVALLIILALGFLTSAQVSPRITFKPIASGQVATCTRRSYSCNEMLDPDCG 433
QY 675 -----SNLIDACTLHAALST-----VS 692
DB 434 FCKMKNSTVLDSSCVPNKASTNEAMGRCENTKFKETEDIFPAVNFCTPPYSWTLLG 493
QY 693 VIVFCCGVMGFGPIPNILCAIEPTTRVRCIGIAICAFETWIGDIIVYSLPVMNAGL 752
DB 494 LILVLFAPMGMPMTVNSEIYPLMARSTGNACSSGIMNIFNVLSLFTLHAELYTY 553
QY 753 AGVPSIYAVVCLISFVYFVFLKVPETKMPLEVITEFP 789
DB 554 YGAFFLYAGFAVGLFTFYGLPETKKGKXLEIESLF 590

RESULT 14
ID GLCP_SYNY3 STANDARD; PRT; 468 AA.
AC P15729;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glucose transport protein.
GN GTR OR GLCP OR SLL0771.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
CX NCBI_TaxId=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90014182; PubMed=2507869;
RA Zhang C.C., Durand M.C., Jeanjean R., Josep F.;
RT "Molecular and genetic analysis of the fructose-glucose transport
RL system in the cyanobacterium Synechocystis PCC6803.";
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=9134660; PubMed=2129397;
RA Schmetterer G.R.;
RT "Sequence conservation among the glucose transporter from the
RL cyanobacterium Synechocystis sp. PCC 6803 and mammalian glucose
RL plant Mol. Biol. 14:697-706 (1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugita M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RL region from map positions 64% to 92% of the genome.";
RN [4]
RX DNA Res. 2:153-166 (1995).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC
DB EMBL, X15988; CA34119.1; -
DB EMBL, X16472; CA34492.1; -
DB EMBL, D64000; BAA10117.1; -
DB PIR, S10014; S10014.
DB InterPro; IPR007114; MFS.
DB InterPro; IPR005828; Sub_transporter.
DB InterPro; IPR005829; Sug_transporter.
DB InterPro; IPR003663; Sugar_transp.
DB Pfam; PF00083; sugar_tr; 1.
DB PRINTS; PR00171; SUGRTNSPORT.
DB TIGRfam; TIGR00879; SP; 1.

```


CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X98622; CA67211.1; -
 DR EMBL; Z98530; CAB1061.1; -
 DR EMBL; Z98532; CAB16718.1; -
 DR PIR; T43400; T43400.
 DR GeneDB; SPombe; SPAC4F8.15; -
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub_transporter.
 DR InterPro; IPR005829; Sub_transporter.
 DR InterPro; IPR003663; Sugar_transp.
 DR Pfam; PF00083; sugar_tr; 1.
 DR PRINTS; PR00171; SUGRTNSPORT.
 DR TIGRfam; TIGR00879; SP; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; FALSE NEG.
 KM Transmembrane; Sugar transport; glycoprotein.
 FT DOMAIN 1 86
 FT TRANSMEM 87 107
 FT DOMAIN 108 129
 FT TRANSMEM 130 150
 FT DOMAIN 151 156
 FT TRANSMEM 157 177
 FT DOMAIN 178 186
 FT TRANSMEM 187 207
 FT DOMAIN 208 215
 FT TRANSMEM 216 236
 FT DOMAIN 237 246
 FT TRANSMEM 247 267
 FT DOMAIN 268 349
 FT TRANSMEM 350 370
 FT DOMAIN 371 376
 FT TRANSMEM 377 397
 FT DOMAIN 398 400
 FT TRANSMEM 401 421
 FT DOMAIN 422 441
 FT TRANSMEM 442 462
 FT DOMAIN 463 486
 FT TRANSMEM 487 507
 FT DOMAIN 508 510
 FT TRANSMEM 511 531
 FT DOMAIN 532 575
 FT CARBOHYD 432 432
 SQ SEQUENCE 575 AA; 62757 MW; 3B7C5E8F86C596AE CRC64;

Query Match 9.9%; Score 404.5; DB 1; Length 575;
 Beat Local Similarity 19.4%; Pred. No. 7.5e-19;
 Matches 153; Conservative 104; Mismatches 192; Indels 341; Gaps 17;

QY 38 ASCSSEPTSDIDLEDK--MSGAVLVAVASIGNLQGMNATTAAYLYI-----87
 DB 64 SWISSSEPHALNDYEDDPVSKWVLAFAAGIGLFGYDGVISGALVVGTSIGGHE 123
 QY 88 ---KKEFQONEPTEVEGLIVSMGLIGATVITTFSGPLSDSIGRRPMLLSILYFFSG 143
 DB 124 LTNGGKEF-----ITSATSGALLGIIAGALADPFGRRKRVIAASIIIVGSI 172
 QY 144 IMLNPNVYVLLARFVDFGIGLAVTVPLVYISFAISEIRGLNLTLPQFSGSGMPLS 203
 DB 173 VGVTAHMLHMTVGRVIGWGVGLSLIPLVLSFAISKIRGLVITVLLITAGQVIA 232
 QY 204 YCMFGMSLS--PSPDRIIMGLVLAISLFFFGLTIFLPSPPRLVSKGRMAEAKVL 260
 DB 233 Y-----GIDPAFPHVNHGMWGLAMVPAFQLFILI-WLPSPRLLVKKERSQEAHYTL 287
 QY 261 QKLRGKDVSGELSLLEGLLEVGGDTISIEYIIIGPATEAADLVTDGKREQITLVGPEEG 320
 DB 288 ARI----- 290

QY 321 QSWIARPSKGPIMLGVSILASRHGSMVNSQVPLMDPIVTLFGSHENMPQAGSMSTL 380
 DB 291 -----YFRAHP----- 296
 QY 381 PPNFGSMFSTVDQAKNQEDEENLHRDDEBYASDAGGDYEDNLHSPLLSRQATGABGK 440
 DB 297 -----YEIK----- 300
 QY 441 DIVHGHRSALSNRQOTLLGEGDGVSTDIGGMQLAMKMSKEGENGRKEGFKRY 500
 DB 301 -----TKLY 304
 QY 501 LHQGVPSGRSGISVLSJGGGVDFEGSEFVHAALVQSALFSKGLAEPMSDAMVHPS 560
 DB 305 LIQEGV-----RDPF----- 314
 QY 561 EVAAGSRMKDL-----FEGVRALLVGVGIQILQAPAGINGVLYTPQILQAGV 612
 DB 315 ---SGSRWQKIVTFPEKELYNPENFRLIACGLQAMQLSGFNSLMYFSTIFEVVG- 369
 QY 613 AVILSKRGLSASASILISLTTLMLPCIGFAMLMNLSGRRLULGCTIPLIASVYL 672
 DB 370 ---FNNPRTATGIIAA--TNEVFTVAFVGI--DFGRRILLTLTWGMIAALIVC 418
 QY 673 VVS-----NLIDGLTALHLLSTVSIVYFCCFVWGFGPIPNILCAEIPPTRVRG 722
 DB 419 AVAHPFLPKDNGNYTSGQSNAMAIIVLISIVIVYASASGLNLP-WQSELEPMSVRG 477
 QY 723 LCIAICAFTEWIGDIIVTVSLPVMNLGAGVESIYAVVCLISFVFVFLKVPETKMP 782
 DB 478 LGTGSTAVMNAAGNLGASFLTLMSEITPTGTALVGGLFCGLMALFCYPDLDVYTI 537
 QY 783 EVITEFPFVVG 792
 DB 538 BEIGELMG 547

Search completed: January 5, 2004, 18:56:02
 Job time : 36 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 18:09:20 ; Search time 95 Seconds
(without alignments)
2173.075 Million cell updates/sec

Title: US-10-051-909-32

Perfect score: 4075

Sequence: 1 IRSGSMIAVQTFPTDLDLR.....PLEVITEFFAVGAKQAANKA 800

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2834	69.5	740	10	Q9FRK0
2	2808	68.9	743	10	Q8GT52
3	2654	65.1	739	10	Q8LPO8
4	2584	63.4	729	10	Q65497
5	2278	55.9	734	10	Q9SY03
6	2275	55.8	734	10	Q9IM67
7	2245	55.1	729	10	Q9SD00
8	2191.5	53.8	753	10	Q8GT51
9	2183	53.6	734	10	Q96290
10	1763	43.3	723	10	Q8H887
11	1763	43.3	723	10	Q9SNK7
12	1270.5	31.2	746	10	Q8LJX7
13	1050	25.8	643	10	Q8H6J2
14	621.5	15.3	664	10	Q9XEVI
15	587.5	14.4	457	16	P96742
16	536	13.2	580	10	Q9C757

17	535.5	13.1	581	10	Q9LKH2	Q9LKH2 mesembryant
18	524.5	12.9	580	10	Q9ZOP6	Q9ZOP6 arabidopsis
19	520.5	12.8	582	10	Q23492	Q23492 arabidopsis
20	519	12.7	581	10	Q9LKH1	Q9LKH1 mesembryant
21	518.5	12.7	469	16	Q97U57	Q97U57 clostridium
22	512.5	12.6	577	10	Q945E5	Q945E5 oryza sativ
23	505	12.4	473	16	Q34718	Q34718 bacillus su
24	500	12.3	509	10	Q8VZB6	Q8VZB6 arabidopsis
25	498	12.2	463	16	Q8EMP5	Q8EMP5 oceanobacil
26	494	12.1	470	10	Q9LLE0	Q9LLE0 solanum tub
27	493.5	12.1	508	10	Q9ZNS0	Q9ZNS0 arabidopsis
28	493	12.1	493	10	Q23213	Q23213 arabidopsis
29	492	12.1	467	16	Q8CQ47	Q8CQ47 staphylococ
30	492	12.0	534	10	Q9LLE1	Q9LLE1 nicotiana t
31	489	12.0	521	10	Q22848	Q22848 arabidopsis
32	488.5	12.0	502	10	Q9FRP7	Q9FRP7 oryza sativ
33	488	12.0	545	10	P93076	P93076 beta vulgar
34	488	12.0	549	10	P93075	P93075 beta vulgar
35	485.5	11.9	539	10	Q9L592	Q9L592 arabidopsis
36	481	11.8	513	10	Q9FOX3	Q9FOX3 apium grave
37	474	11.6	544	10	Q93WT7	Q93WT7 olea europ
38	474	11.6	547	10	Q9SKT9	Q9SKT9 arabidopsis
39	469.5	11.5	515	10	Q9LDB8	Q9LDB8 arabidopsis
40	468	11.5	551	10	Q9LBE2	Q9LBE2 spinacia ol
41	467	11.5	546	10	Q93241	Q93241 arabidopsis
42	465	11.4	419	16	P94493	P94493 bacillus su
43	465	11.4	506	10	Q8W2M8	Q8W2M8 oryza sativ
44	465	11.4	542	10	Q9LID9	Q9LID9 zea mays (m
45	465	11.4	574	10	Q9AUM9	Q9AUM9 oryza sativ

ALIGNMENTS

RESULT 1
Q9FRK0 PRELIMINARY; PRT; 740 AA.
ID Q9FRK0; AC Q9FRK0; DT 01-MAR-2001 (TRMBLrel. 16, Created)
DT 01-MAR-2001 (TRMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TRMBLrel. 23, Last annotation update)
DE Putative sugar transporter.
OS OSUNB0064P21.3.
GN Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell C.R., Yuan Q., Mofatt K.S., Hill J.N., Burr P.C., Hsiao J.,
RA Zismann V., Pai G., Bowman C.D., Fujii C.Y., Vanaken S.E.,
RA Bowman C.D., Craven B., Utterback T.R., Khalak H., Feldblyum T.V.,
RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 10 BAC OSUNB0064P21 genomic sequence.";
RL Submitted (AUG-2001) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC EMBL; AC073166; AAG46115.1; -
CC Gramene; Q9FRK0; -
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transp.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar tr; 1.
DR PRINTS; PR00171; SUGTRANSPORT.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 740 AA; 78899 MW; DAC003E0A4980C46 CRC64;

Query Match 69.5%; Score 2834; DB 10; Length 740;
 Best Local Similarity 74.5%; Pred. No. 2.9e-184;
 Matches 553; Conservative 75; Mismatches 108; Indels 6; Gaps 5;

56 MSGAVLVAIVASIGULQGMNATTAATVLYTKKEFOLQNEPTVBGLVSMGLGATVYT 115
 1 MAGAVLVVAIVASIGULQGMNATTAATVLYTKKEFOLQNEPTVBGLVSMGLGATVYT 60

116 TFSGLSDSISGRPMILISILYFSGGLIMLSPNVTYVLLARFVDFGIGLAVTLVPLY 175
 61 TFSGLVADSFGRRPMLIASVLYFVSGVLMAPVYVLLARLIDGFGIGLAVTLVPLY 120

176 ISEIAPSEIRGLANTLPQFSGSGMFLSYCVFSGSLSPSPDMRIMLGVAIPSLFFGL 235
 121 ISEIAPTDIRGLANTLPQFSGSGMFLSYCVFSGSLMPQDMRIMLGVAIPSLFFGL 180

236 TIFYPESPRLVSGRNAEAKVYOKLRGKODVSGELSLLEGLEVGDDTSIEEYITGP 295
 181 TIFYPESPRLVSGRNAEAKVYOKLRGKODVSGELSLLEGLEVGDDTSIEEYITGP 240

296 ATEAADLVTDGDKQITLYGPEGQSWIARPSKGMIMGSLVSLASRHSVMNQSVPLM 355
 241 DDELADDEGLAP-DPEKIKLYGPEEGISVYARVHGQALGSLALISRHGSMVSGKPLV 299

356 DEIVTLFSGVHNMPOAGSMRSTLPNFGSMFSTVDQAKNQBQDEBNLHRDDEYASD 415
 300 DPVVTLYFSGVHNMPOAGSMRSTLPNFGSMFSTVAEQQAKADWDAS-QREGEDYSD 358

416 GAGGYEDNHSPLSRQATGAEKDI-VHGHRSALSMMRQITLLGEGDGVSTDIIG 474
 359 HGGDDIEDLSPLSRQATVGEKEIAPHGSIKAVG--RSSSLMOGGEAVSSMGIG 416

475 GMLAMKSEKENGKRGKGFYVLYHQEGVPSGRGSIYSLPGGQGVFEGSEFVHAA 534
 417 GMLAMKTEREGADGKRGKGFYVLYHQEGVPSGRGSIYSLPGGQGVFEGSEFVHAA 475

535 LVSGSALFSGKLAEPMSDAMVHPSEVAAKSRMKDLFEPGVRRALLVGVGIQILQQA 594
 476 LVSGSALFSGKLAEPMSDAMVHPSEVAAKSRMKDLFEPGVRRALLVGVGIQILQQA 535

595 GINGVLYTTPQILEAGVAVILSKFGLSSASASIIISLTITLMLPCIGFAMLMDSGR 654
 536 GINGVLYTTPQILEAGVAVILSKFGLSSASASIIISLTITLMLPSIGIARMLDMSGR 595

655 RFLILGTIPILIASIVILVSNLIDGLTAHALSTVSIVYFCFWMGFGPIPNILCAE 714
 596 RFLILATIPILIALILVILVNLIDVGMVHASLSTVSIVLYFCFWMGFGPIPNILCAE 655

715 IFPTVRVGLCIAICAFTEWIGDIIVTYSIPVNLNAGLAGVSIYAVVCLISFVFPFLV 774
 656 IFPTVRVGLCIAICAFTEWIGDIIVTYSIPVNLNAGLAGVSIYAVVCLISFVFPFLV 715

775 PETKGMPELVITEFPAVGAQA 796
 716 PETKGMPELVITEFPAVGAQA 737

RESULT 2
 08GTS2 PRELIMINARY; PRT; 743 AA.
 AC 08GTS2;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Hexose transporter.
 GN Strp1.
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticaceae; Hordeum.
 OC NCB1_taxonomy=4513;
 RP [1]
 RN SEQUENCE FROM N.A.

RC STRAIN=cv. Barke; TISSUE=Developing caryopses;
 RA Meschke W., Pantz R., Gubatz S., Wang Q., Radchuk R., Weber H.,
 RA Wobus U.;
 RT "The role of invertases and hexose transporters in controlling sugar
 RT ratios in maternal and filial tissues of barley caryopses during early
 RT development";
 DR Submitted (DEC-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AJ534445; CAD58958.1; -
 SQ SEQUENCE 743 AA; 79339 MW; 0169B6BD06007D5 CAC64;

Query Match 68.9%; Score 2808; DB 10; Length 743;
 Best Local Similarity 73.1%; Pred. No. 1.7e-182;
 Matches 543; Conservative 84; Mismatches 104; Indels 12; Gaps 6;

57 SGAVLVVAIVASIGULQGMNATTAATVLYTKKEFOLQNEPTVBGLVSMGLGATVYT 116
 3 SGAVLVVAIVASIGULQGMNATTAATVLYTKKEFOLQNEPTVBGLVSMGLGATVYT 62

117 FSGPLSDSISGRPMILISILYFSGGLIMLSPNVTYVLLARFVDFGIGLAVTLVPLY 176
 63 FSGVAVADVGRPMILIASVLYFVSGVLMAPVYVLLARLIDGFGIGLAVTLVPLY 122

177 SEIAPSEIRGLANTLPQFSGSGMFLSYCVFSGSLSPSPDMRIMLGVAIPSLFFGL 236
 123 SETAPTDIRGLANTLPQFSGSGMFLSYCVFSGSLMPQDMRIMLGVAIPSLFFGL 182

237 IEFYPESPRLVSGRNAEAKVYOKLRGKODVSGELSLLEGLEVGDDTSIEEYITGP 296
 183 IEFYPESPRLVSGRNAEAKVYOKLRGKODVSGELSLLEGLEVGDDTSIEEYITGP 242

297 TEAADLVTDGDKQITLYGPEGQSWIARPSKGMIMGSLVSLASRHSVMNQSVPLM 354
 243 DELADDDGLAP-DQKIKLYGPEEGISVYARVHGQALGSLALISRHGSMVSGKPLV 301

355 MDPIVTLFSGVHNMPOAGSMRSTLPNFGSMFSTVDQAKNQBQDEBNLHRDDEYAS 414
 302 VDPVVTLYFSGVHNMPOAGSMRSTLPNFGSMFSTVAEQQAKADWDAS-HRDEYAS 360

415 DGAGGYEDNHSPLSRQATGAEKDI-VHGHRSALSMMRQITLLGEGDGVSTDIIG 472
 361 DGAGGYEDNHSPLSRQATVGEKEIAPHGSIKAVG--RSSSLMOGGEAVSSMGIG 415

473 GMLAMKSEKENGKRGKGFYVLYHQEGVPSGRGSIYSLPGGQGVFEGSEFVHAA 532
 416 GMLAMKTEREGADGKRGKGFYVLYHQEGVPSGRGSIYSLPGGQGVFEGSEFVHAA 474

535 LVSGSALFSGKLAEPMSDAMVHPSEVAAKSRMKDLFEPGVRRALLVGVGIQILQQA 592
 475 LVSGSALFSGKLAEPMSDAMVHPSEVAAKSRMKDLFEPGVRRALLVGVGIQILQQA 534

593 GINGVLYTTPQILEAGVAVILSKFGLSSASASIIISLTITLMLPCIGFAMLMDSGR 652
 535 GINGVLYTTPQILEAGVAVILSKFGLSSASASIIISLTITLMLPSIGIARMLDMSGR 594

653 RFLILGTIPILIASIVILVSNLIDGLTAHALSTVSIVYFCFWMGFGPIPNILCAE 712
 595 RFLILSTIPILIALILVILVNLIDVGMVHASLSTVSIVLYFCFWMGFGPIPNILCAE 654

713 IFPTVRVGLCIAICAFTEWIGDIIVTYSIPVNLNAGLAGVSIYAVVCLISFVFPFLV 772
 655 IFPTVRVGLCIAICAFTEWIGDIIVTYSIPVNLNAGLAGVSIYAVVCLISFVFPFLV 714

773 KVPETKGMPELVITEFPAVGAQA 795
 715 KVPETKGMPELVITEFPAVGAQA 737

RESULT 3
 08LP08 PRELIMINARY; PRT; 739 AA.
 AC 08LP08;
 DT 01-OCT-2002 (Tremblrel. 22, Created)
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE AT4G35300/F23E12.140.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eucotylidae II; Brassicales; Brassicaceae; Arabidopsids.
 CC NCBI_TaxID=3702;
 CC [1]
 RP SEQUENCE FROM N.A.
 RA Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J.,
 RA Bowser L., Cammici P., Chang B., Dale J.M., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Saou M.,
 RA Seki M., Southwick A., Tang C.C., Tortum M., Wu H.C., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinzaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RA "Arabidopsids cDNA clones";
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC EMBL, AY094465; AAA19835.1; -.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub_transporter.
 DR InterPro; IPR003663; Sugar_transp.
 DR InterPro; IPR005829; Sug_transporter.
 DR Pfam; PF00083; sugar_tr.2.
 DR PRINTS; PR00171; SUGTRANSPORT.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Sugar transporter; Transmembrane; Transport.
 KW SEQUENCE 739 AA; 79725 MW; 5D67C6D2DE7563FD CRC64;

Query Match 65.1%; Score 2654; DB 10; Length 739;
 Best Local Similarity 70.0%; Pred. No. 5,1e-172;
 Matches 521; Conservative 90; Mismatches 123; Indels 10; Gaps 7;

QY 56 MSGATLVAVASIGMLTQGMNATTAALVLYIKKFPOLNERTVYGLVMSLIGATVY 115
 DB 1 MSGATLVAVAAVGLTQGMNATTAAGAVLYIKKFEFNSNSVGLVAVMSLIGATIT 60
 QY 116 TFSGSLVSGRRPMLTSSILYFSGIMLSPVNYVLLARFPDVGIGIAVTVPLV 175
 DB 61 TCSGVAALWLRGRPMILSSILYFVGLVLMSPVNYVLLARLLDGVGVGLVTVPLV 120
 QY 176 ISEIAPSEIRGLNLTLPQFSGGMLSCVNFVGNLSPPSPDRIMLGVALPSLFFGL 235
 DB 121 ISEIAPSEIRGLNLTLPQFSGGMLSCVNFVGNLSPPSPDRIMLGVALPSLFFGL 180
 QY 236 TIFVLPSEPRMIVSGMAEAKVYOKLRKDVSGELSLLEGLVEVGDTISIEVYIGP 295
 DB 181 TIFVLPSEPRMIVSGMAEAKVYOKLRKDVSGELSLLEGLVEVGDTISIEVYIGP 240
 QY 296 ATEAADLVLTQDQKQITLYGPEEGSWIARPSKPIMLGVALSLASHGS-MVAVQSVL 354
 DB 241 ADEVTDDIADVQKQITLYGPEEGSWIARPSKPIMLGVALSLASHGS-MVAVQSVL 296
 QY 355 MDPIVTLFSGVHNNPQAGSGNRSTLPNFVSGMFTVQHAKEQWDEENILHRDDEYAS 414
 DB 297 IDPLVTLFSGVYHEKMPDT-GSWRSALPFPFSGMFTVQHAKEQWDEENILHRDDEYAS 355
 QY 415 DGAGGDIYEDNLHSLPLSRQATAGKQDVHGHRSALSMMRQTLIGSGGDSVSTDIG 474
 DB 356 D-HGDSDDDLHSLPLSRQATAGKQDVHGHRSALSMMRQTLIGSGGDSVSTDIG 413
 QY 475 GMLQAKKSEKGEKGRKGFVYLYHOGVPGVSRGIVSLPGSGDVFESEFVHAA 534
 DB 414 GMLQAKKSEKGEKGRKGFVYLYHOGVPGVSRGIVSLPGSGDVFESEFVHAA 472
 QY 535 LVSGALVPSKGLAEPFMSDAVAVHPSSEVAVAKSRWKDLEPEPVRALLVGVGIQLQOP 594
 DB 473 LVSGALVPSKGLAEPFMSDAVAVHPSSEVAVAKSRWKDLEPEPVRALLVGVGIQLQOP 531

QY 595 GINGVLYTTPQILEAGVAVILSKRGLSSASATILISLTLLMPCIGPAMLLMDLSGR 654
 DB 532 GINGVLYTTPQILEAGVAVILSKRGLSSASATILISLTLLMPCIGPAMLLMDLSGR 591
 QY 655 RELLGTPILIASVILVAVNLDGLTAHALSTVSVIYVFCFVWGFGPIPNILCAE 714
 DB 592 RELLGTPILIASVILVAVNLDGLTAHALSTVSVIYVFCFVWGFGPIPNILCAE 651
 QY 715 IFFPRVRLCIAICFTWIGDIYVTSLPVMAAIGAGFSTIYAVVCLISFVVFELV 774
 DB 652 IFFPRVRLCIAICFTWIGDIYVTSLPVMAAIGAGFSTIYAVVCLISFVVFELV 711
 QY 775 PETKMPLEVTTEPFAVAKQAA 798
 DB 712 PETKMPLEVTTEPFAVAKQAA 735

RESULT 4

ID 065497 PRELIMINARY; PRT; 729 AA.
 AC 065497;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Putative sugar transporter protein (Monosaccharide sensing protein 2).
 GN F23E12.140 OR AT4G35300 OR MSSP2.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eucotylidae II; Brassicales; Brassicaceae; Arabidopsids.
 CC NCBI_TaxID=3702;
 CC [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,
 RA Hohnselt J., Jesse T., Heijnen L., Vos P., Mewes H.W., Mayer K.F.X.,
 RA Schueller C.;
 RL Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,
 RA Mewes H.W., Lemcke K., Mayer K.F.X.;
 RL Submitted (Mar-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (Mar-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Stamme C., Tjaden J., Trentmann O., Emmerlich V., Neuhaus E.;
 RT "A novel family of plant monosaccharide transporters is involved in a
 RT new type of eukaryotic sugar sensing";
 RT Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC EMBL, AL022604; CAAL8739.1; -.
 CC EMBL, AL161587; CAB80247.1; -.
 CC EMBL, AJ532570; CAB80247.1; -.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub_transporter.
 DR InterPro; IPR003663; Sugar_transp.
 DR InterPro; IPR005829; Sug_transporter.
 DR Pfam; PF00083; sugar_tr.1.
 DR PRINTS; PR00171; SUGTRANSPORT.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Sugar transporter; Transmembrane; Transport.
 KW SEQUENCE 729 AA; 78524 MW; 81A35DDF8725411E CRC64;

Query Match 63.4%; Score 2584; DB 10; Length 729;
 Best Local Similarity 68.8%; Pred. No. 2.9e-167;

Matches 512; Conservative 89; Mismatches 123; Indels 20; Gaps 8;

QY 56 MSGATLVIVASISGNLQGMNATTAATAVLYIKKEPOLONEPTVEGLIVMSLIGATIVT 115
 DB 1 MSGATLVIVAAVAVGNLQGMNATTAAGALVYIKKEZNLSENSVSGVLIVAMSLIGATITLT 60
 QY 116 TFSGPIVLSIGRRPMLILSSILYFSGILMISPNVYVLLARFVDFGFIGLAVTLVPLV 175
 DB 61 TCGGVAAVWMLGRPMILILSSILYFVGLVLMSPVYVLLARLDRGLDGVGLVTLVPLV 120
 QY 176 ISEIAPSEIRGLANTLPQFSGSGMPLSYCMVFGMSLSPSPWIMLGVALIPSLFFPGL 235
 DB 121 ISEIAPSEIRGLANTLPQFSGSGMPLSYCMVFGMSLSPSPWIMLGVALIPSLFFPGL 180
 QY 236 TIFVLPESPRMIVSGRMAEAKVLOKRGKDVSGELSLLEGVEGDDTSIEEYITGP 295
 DB 181 TIFVLPESPRMIVSGRMAEAKVLOKRGKDVSGELSLLEGVEGDDTSIEEYITGP 240
 QY 296 ATEAADLVTDGDKQITLYGPEGQSWIARPSKPIMLGVLISLRHGS-MVNSQVPL 354
 DB 241 ADEVTDHDIADVNDKQITLYGABEGLSWARPVK-----GSTMSVLSRHGSTMRRQSL 296
 QY 355 MDPIVTLFGSVHNNPQAGSGNRSTLPFNFGSMFSVITDQAHAKNEDENLRHDEEYAS 414
 DB 297 IDPLVTLFGSVHEKKPDT-GSMRSALFPHFGSMFSGVGNQPHEDMDENLVGEGEDYPS 355
 QY 415 DGAGDVEDNHSPLLSRQATGAEKDIYHGHRSALSMRQITLLEGEGDVSSTIDCG 474
 DB 356 D-HGDDSDDDLSPLISRQITSMK-KDMPHTAHGTLSTFRHGSQVQAGQEGAGSGMGIG 413
 QY 475 GMLMKNSKEKENGKREKGFERYLHQEGVPSNRGSIVSLPGGADVFESEBFVHA 534
 DB 414 GMLVAMKTERDESGOKE-----EGFPGSRGSIVSLPGGDDGFE-ADVQVQSA 462
 QY 535 LVSGALFSKGLAEPMSDAMVHSEVAAKSRMDLFEPEVRALLVGVIGIQLQFPA 594
 DB 463 LVSGALFSKGLAEPMSDAMVHSEVAAKSRMDLFEPEVRALLVGVIGIQLQFPA 521
 QY 595 GINGLYTTPQILBOAGVAVILSKFGLSASASLISLITLMLPCIGFALMLDLSGR 654
 DB 522 GINGLYTTPQILBOAGVAVILSKFGLSASASLISLITLMLPALAVALRMLDLSGR 581
 QY 655 RPLILGTPILILASVILVAVNSLIDGLTAHALSTVSIVYFCEFGVGFPIPNILCAE 714
 DB 582 RPLILGTPILILASVILVAVNSLIDGLTAHALSTVSIVYFCEFGVGFPIPNILCAE 641
 QY 715 IFFTRVRCICAIACFTFWIGDIIVTYSLPVLMALIGLAVFSIYAVVCLISFVFLKV 774
 DB 642 IFFTRVRCICAIACFTFWIGDIIVTYSLPVLMALIGLAVFSIYAVVCLISFVFLKV 701
 QY 775 PETKMPLEVITEFPAVGAQA 798
 DB 702 PETKMPLEVITEFPAVGAQA 725

RESULT 5
 Q9SYO3 PRELIMINARY; PRT; 734 AA.
 AC Q9SYO3;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Sugar transporter.
 GN P9H16.18.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucotids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_Taxid=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Federjpiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,

RA Alfafi H., Araujo R., Huizer L., Rowley D., Buehler E., Dunn P.,
 RA Gonzalez A., Kremenetskaia I., Kim C., Ienz C., Li J., Liu S.,
 RA Lucos S., Schwartz J., Shim P., Toriumi M., Vysotskaia V.S.,
 RA Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;
 RL Submitted (May-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL; AC007369; AAD30608.1; -.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub_transporter.
 DR InterPro; IPR003663; Sugar_transporter.
 DR InterPro; IPR005829; Sug_transporter.
 DR Pfam; PF00083; sugar_tr; 1.
 DR PRINTS; PR00171; SUGRTRANSPORT.
 DR PROSITE; PS00550; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 DR Sugar transporter; Transmembrane; Transport.
 SQ SEQUENCE 734 AA; 79426 MW; AB027864F850259C CRC64;

Query Match 55.9%; Score 2278; DB 10; Length 734;
 Best Local Similarity 61.2%; Pred. No. 2e-146;
 Matches 462; Conservative 112; Mismatches 145; Indels 36; Gaps 14;

QY 56 MSGATLVIVASISGNLQGMNATTAATAVLYIKKEPOLONEPT-VEGLIVMSLIGATIV 114
 DB 1 MKGATLVLAATIGNFLQGMNATTAAGAMVYINDNL---PTVQGLVAMSLIGATVI 57
 QY 115 TFSGPIVLSIGRRPMLILSSILYFSGILMISPNVYVLLARFVDFGFIGLAVTLVPL 174
 DB 58 TTGCGPIVDMVGRPMILILSSVMTFVCGILVMSPNVYVLCFARLNFAGLAVTLVPLV 117
 QY 175 ISEIAPSEIRGLANTLPQFSGSGMPLSYCMVFGMSLSPSPWIMLGVALIPSLFFPFG 234
 DB 118 ISEIAPSEIRGLANTLPQFSGSGMPLSYCMVFGMSLSPSPWIMLGVALIPSLFFPFG 177
 QY 235 LTIPTLPSPMLVSKGRMAEAKVLOKRGKDVSGELSLLEGVEGDDTSIEEYITGP 294
 DB 178 LTIPTLPSPMLVSKGRMAEAKVLOKRGKDVSGELSLLEGVEGDDTSIEEYITGP 237
 QY 295 PATEAADLVTDGDKQ- QITLYGPEGQSWIARPSKPIMLGVLISLRHGS-MVNSQV 353
 DB 228 LEDHEGDDTLTVEDEGMRLYGTENQSYLARVPBQ---NSSIGLRSRGSILANQMI 294
 QY 354 LMDPIVTLFGSVHNNPQAGSGNRSTLPFNFGSMFSVITDQ-HAKNEWD---EENLRD 408
 DB 295 LMDPIVTLFGSVHNNPQAGSGNRSTLPFNFGSMFSVITDQ-HAKNEWD---EENLRD 354
 QY 409 DEEYASDAGGDYED---NLHSPLLSRQATGAEKDIYHGHRSALSMRQITLLEGEGD 465
 DB 355 NDDYATDAGADDDSDNDLSPLMSRQITSMK-KDMIPHTSGTSLSMRHSITLMQ-GN 412
 QY 466 GVSSTIDGGQMLAMKNSKEKENGKREKGFERYLHQEGVPSNRGSIVSLPGGADVFE 525
 DB 413 GESSWIGCGVHMGVRYENDE-----YKRYLTKEDGAE-SRRGSIIISIPGPD--G 460
 QY 526 GSEFVHAALVYSQALFSKGLAEPMSDAMVHSEVAAKSRMDLFEPEVRALLVGV 585
 DB 461 GSEFVHAALVYSQALFSKGLAEPMSDAMVHSEVAAKSRMDLFEPEVRALLVGV 515
 QY 586 GIQLIQPAGINGVLYTTPQILBOAGVAVILSKFGLSASASLISLITLMLPCIGFA 645
 DB 516 GIQLIQPAGINGVLYTTPQILBOAGVAVILSKFGLSASASLISLITLMLPCIGFA 575
 QY 646 MLMDLSGRFLLGTPIILILASVILVAVNSLIDGLTAHALSTVSIVYFCEFGVFG 705
 DB 576 MLMDLSGRFLLGTPIILILASVILVAVNSLIDGLTAHALSTVSIVYFCEFGVFG 635
 QY 706 PIPNILCEIPTTRVRCICAIACFTFWIGDIIVTYSLPVLMALIGLAVFSIYAVVCL 765
 DB 636 PIPNILCEIPTTRVRCICAIACFTFWIGDIIVTYSLPVLMALIGLAVFSIYAVVCL 695
 QY 766 SFVFLVLPETKMPLEVITEFPAVGAQA 800

Db 696 SMIFVYMKVPETKCMPLVITDYFAFGA-QAQAASA 729

RESULT 6

Q9LM67 PRELIMINARY; PRT; 734 AA.

AC Q9LM67; 01-OCT-2000 (Tremblrel. 15, last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, last annotation update)
 DE F2D10.36.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosida II; Brassicales; Brassicaceae; Arabidopsia.
 NCBI_TaxID=3702;
 RX [1]
 RN
 RP SEQUENCE FROM N.A.
 RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
 RA Kim C., Altafi H., Bei Q., Chin C., Chio J., Choi E., Conn L.,
 RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
 RA Lenz C., Li J., Liu A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
 RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
 RA Ecker J.R.;
 RT "Genomic sequence for Arabidopsis thaliana BAC F2D10 from chromosome
 I.";
 RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL; AC069251; AAF0627.1; -
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; SUG_TRANSPORTER.
 DR InterPro; IPR003663; Sugar_transp.
 DR InterPro; IPR005829; Sug_transporter.
 DR Pfam; PF00083; sugar_tr; 1.
 DR PRINTS; PR00171; SUGTRNSPORT.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Sugar transport; Transmembrane; Transport.
 SQ SEQUENCE 734 AA; 79485 MW; 0232F8BF850259F CRC64;

Query Match 55.8%; Score 2275; DB 10; Length 734;
 Best Local Similarity 61.2%; Pred. No. 3, 1e-146;
 Matches 462; Conservative 111; Mismatches 146; Indels 36; Gaps 14;
 Oy 56 MSGAVLVIVASIGLQGMNATIAAVALYIKKEFOLQNEPT-VEGLIVSMILGATIV 114
 Db 1 MKGATLVLAATIGNFLQGMNATIAAGAVYINKLNL---PTVQGLVVAASLIGATVI 57
 Oy 115 TTFSGLPDSIGRRPMLISLILYFSGILMLSDNVYVLLARFVDFSGIGLAVTVPL 174
 Db 58 TTSGSPISDMGRMRMLISLIVYVCGILMMSNVYVLCARLLNGGGLATVLPV 117
 Oy 175 YISEIAPEIRGLNTLPQFSSGGMFLSYCVFSGMSLSPSPDWIMGLVTAIPSLFFFG 234
 Db 118 YISEIAPEIRGLNTLPQFSSGGMFLSYCVFSGMSLSPSPDWIMGLVTAIPSLFFFG 177
 Oy 235 LTIPLPSPRLVSKGMAEKAYVLOKRGDVSGLSLLBGLVGVGDSISEYIIG 294
 Db 178 LTIPLPSPRLVSKGMAEKAYVLOKRGDVSGLSLLBGLVGVGDSISEYIIG 237
 Oy 295 PATEAADDVLTGDXE-QITLYPREGQSWIARPEKGIIMLSVLSLRHSGMNVQSP 353
 Db 238 LEDHGDDTLTETVDGQWRLYGTGHNSGYLARVPYEQ---NSISGLSRHSGSLANQSMI 294
 Oy 354 LMDPLVTLFSGVHEMMPQAGSGMSTLPPNFSMSVTDQ--HAKNEQD---EENLRD 408
 Db 295 LMDPLVTLFSGVHEMMPQAGSGMSTLPPNFSMSVTDQ--HAKNEQD---EENLRD 354
 Oy 409 DEEYASDAGGDYED--NLHSPILSRQATGAEGKDIYVHGHRGSLASWRQTLTGEGGD 465

Db 355 NDDVATDDGAGDDDDSDNDLRSPLMSRQTTSM-D-KDMLPHPTSGSTLSMRHSTLMQ-GN 412
 Oy 466 GVSSTDIQGGQOLAMKSKSEKENGKRGKRYLTHOEVGPGSRGSIVSLPGGQVFE 525
 Db 413 GESSWGIGGGWHMOWRYENDE-----YKRYYLEKDGAE-SRRGSIIISPGPD--G 460
 Oy 526 GSEFVHAALVSGSALSKGLAEPMSDAAVHSEVAAKSRMKDLEPGVRALVGV 585
 Db 461 GGSYTHASALVSRKSVLGPKS-----VHGSANVPEKTAASGPLMSALIEPGVKALVGV 515
 Oy 586 GIQIIQOPAGINGVLYTTPQILEQAGVAVILSKFGLSASASIISSLTLLMLPCIGFA 645
 Db 516 GIQIIQOPSGINGVLYTTPQILEQAGVAVILSKFGLSASASIISSLTLLMLPAIVVA 575
 Oy 646 MLMDLSGRFLLIGTIPILIASIYIVVSNLIDGLTAHALSTVSIVYVFCFPWFG 705
 Db 576 MRLMDVSGRRSLMTIPVLIVSLVAVISELHISKVNAALSTGCVVLYCFEFGWYG 635
 Oy 706 PIPNLCARPPTRVRCICAFPTWIGIITVTSIPVWMLNIGLAVSIVAVCTI 765
 Db 636 PIPNLCARPPTRVRCICAFPTWIGIITVTSIPVWMLNIGLAVSIVAVCTI 695
 Oy 766 SFVFPVFLKVPETKCMPLVITDYFAFGA-QAQAASA 800
 Db 696 SMIFVYMKVPETKCMPLVITDYFAFGA-QAQAASA 729

RESULT 7

Q9SD00 PRELIMINARY; PRT; 729 AA.

AC Q9SD00; 01-MAY-2000 (Tremblrel. 13, last sequence update)
 DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, last annotation update)
 DE Sugar transporter-like protein (Monosaccharide sensing protein
 3).
 GN F26013.130 OR MS8P3.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosida II; Brassicales; Brassicaceae; Arabidopsia.
 NCBI_TaxID=3702;
 RX [1]
 RN
 RP SEQUENCE FROM N.A.
 RA Delaney M., Berger C., Cooke R., Grellert F., Laurie M., Mewes H.W.,
 RA Lemcke K., Meyer K.F.X., Quetier F., Salanoubat M.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Stamm C., Tjaden J., Trentmann O., Emmerlich V., Neuhaus E.;
 RT "A novel family of plant monosaccharide transporters is involved in a
 new type of eukaryotic sugar sensing.";
 RT Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL; AL133452; CAB63103.1; -
 DR EMBL; AJ532571; CAB58693.1; -
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub_transporter.
 DR InterPro; IPR003663; Sugar_transp.
 DR InterPro; IPR005829; Sug_transporter.
 DR Pfam; PF00083; sugar_tr; 1.
 DR PRINTS; PR00171; SUGTRNSPORT.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Sugar transport; Transmembrane; Transport.
 SQ SEQUENCE 729 AA; 78828 MW; 241A9415DB43FCCC CRC64;

Query Match 55.1%; Score 2245; DB 10; Length 729;
 Best Local Similarity 60.4%; Pred. No. 3 4e-144;
 Matches 457; Conservative 110; Mismatches 144; Indels 46; Gaps 14;

56 MSGAVLVAIVASIGNLQGMNATIAAVALYIKKEFQLONEPTVEGLIVSMIGATVY 115
 1 MRSVVLVLAIAAGNMLQGMNATIAAGAVIYIKKEFHEKEPKRGLIVAMSLIGATVIT 60
 116 TFSGLPSIGRRPMLIISIIYFSGLMSPNVTYLLARFVDPGIGLAATVLPY 175
 61 TFSGVSDVGRKSMILISSVLYFSSIVFMSPNVTYLLARFVDPGIGLAATVLPY 120
 176 ISEIAPSEIRGLINTLPQFSGSGMFLSYCMVFGMSLSPSPMRIMLVIAIPSPFGL 235
 121 ISEIAPSEIRGLINTLPQFSGSGMFLSYCMVFGMSLSPSPMRIMLVIAIPSPFGL 180
 236 TIFVLPESPRLVSGRMAEAKVLOKLRGKDVSGEISLLEGLVEGDTISEEYITGP 295
 181 AAFELPESPRLVSGRMAEAKVLOKLRGKDVSGEISLLEGLVEGDTISEEYITGP 240
 226 ATEAADLVTDGDKQIITLYGPEGQSWTARPSKQIMLGVTSLASRHSVNVOSVPLM 355
 241 DNEENEGNELPRKQIILKYGEDQSWMAKPVKQ---SSLALASRGSVLPFGSLM 296
 356 DEIVTLFSGVHEHMF--QAGSMRSTLPNFGSMFSTVDQAKNOMPENLHRDDEYA 413
 297 DPLVTLFSGIHENLSENNASRSMLFPNMGSIILAMGR--QESQWDE--RNED-- 349
 414 SDGAGDYEDNLHSPILSRQATGAEGKDIIVHGHKRSALSMSRQTL-LGEGGDVSSTDI 472
 350 ---SSDQDENLNSPLSPQTT--EPDD--YHQRVTGTHRRQSSLFMANVGETATATSI 401
 413 GGGNLLARKMSKEENGKRGK--EGGRKRYTLHOE-----GVPSRRGSIVSL-RGGDV 523
 402 GGGNLLARKMSKEENGKRGK--EGGRKRYTLHOE-----GVPSRRGSIVSL-RGGDV 461
 524 FECSFVHAALVSGSALFSGKLAEPMSDAMVHPSEVAAGSKSMKDLFEPGVRRALV 583
 482 DQVNGYVQAALVSGSALFSGKLAEPMSDAMVHPSEVAAGSKSMKDLFEPGVRRALV 514
 584 GVGIIILQOPAGINCVLYTTPQILEQAGVAVILSKRGLSSASASIISSLTLLMLPCIG 643
 515 GVGIIILQOPAGINCVLYTTPQILEQAGVAVILSKRGLSSASASIISSLTLLMLPCIG 573
 644 FAMLMDLSGRFLLGTPIILIASVILVNSLIDGLTAAALSTYSVIVYPCCFPMG 703
 574 ---LVMSKSLMLSTPIILIASVILVNSLIDGLTAAALSTYSVIVYPCCFPMG 626
 704 FGPIPNILCAEIPFTRVGLCIAICAFPMIGDIIVTYSLPVMLNAGIAGVFSIYAVVC 763
 627 FGPIPNILCAEIPFTRVGLCIAICAFPMIGDIIVTYSLPVMLNAGIAGVFSIYAVVC 686
 764 LISFVFLKVPETKMPLEVTTEPFAVGAQAQA 800
 687 AVAMVFLKVPETKMPLEVTTEPFAVGAQAQA 723

RESULT 8

Q8GT51 PRELIMINARY; PRT; 753 AA.
 AC O8GT51;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, last annotation update)
 DE Sugar transporter.
 GN STR2.
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Hordeum.
 OC NCBI_TaxID=4513;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=Cv. Barke; TISSUE=Developing caryopsis;
 RA Weehee W., Panitz R., Gubatz S., Wang Q., Radchuk R., Weber H.,
 RA Wobus U.;
 RT "The role of invertases and hexose transporters in controlling sugar
 RT ratios in maternal and filial tissues of barley caryopses during early
 RT development."
 RT Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ534446; CAD58959.1; -
 SQ SEQUENCE 753 AA; 81489 MW; A533B2B31922A9B CRC64;

Query Match 53.8%; Score 2191.5; DB 10; Length 753;
 Best Local Similarity 66.2%; Pred. No. 1.6e-140;
 Matches 440; Conservative 76; Mismatches 132; Indels 17; Gaps 9;

137 LYPFSGLMSPNVTYLLARFVDPGIGLAATVLPYISIRAPSEIRGLINTLPQFSG 196
 65 LVPFASNVMLAPNVTYLLARFVDPGIGLAATVLPYISIRAPSEIRGLINTLPQFSG 124
 197 SGGMFLSYCMVFGMSLSPSPMRIMLVIAIPSPFGLTIFVLPESPRLVSGRMAEA 256
 125 SGGMFLSYCMVFGMSLSPSPMRIMLVIAIPSPFGLTIFVLPESPRLVSGRMAEA 184
 257 KKVLOKLRGKDVSGEISLLEGLVEGDTISEEYITGPATEAADLVTDGDKQIITLYG 316
 185 KKVLOKLRGKDVSGEISLLEGLVEGDTISEEYITGPATEAADLVTDGDKQIITLYG 242
 317 PEEGQSWTARPSK--GPIMLGSVTLASRHSVNVOSVPLMDPIVTLFSGVHEHMFQAG 374
 243 AEEGQSWTARPSK--GPIMLGSVTLASRHSVNVOSVPLMDPIVTLFSGVHEHMFQAG 302
 375 SMRSTLPNFGSMFSTVDQAKNOMPENLHRDDEYASAGAGDYEDNLHSPILSRQ 434
 303 SMRSTLPNFGSMFSTVDQAKNOMPENLHRDDEYASAGAGDYEDNLHSPILSRQ 361
 435 TGAEGKDIIV--HHGRGSALSMSRQTLIGEGGDVSSTDIGGQVLAMKMSKEENGKRGK 492
 362 TSVGKEIAPHSGIMGVSESSMQ-----GGDAVSSMGICGGQVLAMKMSKEENGKRGK 416
 493 EGGRKRYTLHOEYVPSRRGSIVSLPGGQDVFECSFVHAALVSGSALFSGKLAEPMS 552
 417 EGGRKRYTLHOEYVPSRRGSIVSLPGGQDVFECSFVHAALVSGSALFSGKLAEPMS 475
 553 DAAMVHSEVAAGSKSMKDLFEPGVRRALVGVGILOOPAGINCVLYTTPQILEQAGV 612
 476 GPAMVHSEVAAGSKSMKDLFEPGVRRALVGVGILOOPAGINCVLYTTPQILEQAGV 535
 613 AVILSKRGLSSASASIISSLTLLMLPCIGFAMLMDLSGRFLLGTPIILIASVITL 672
 536 GILLSNIGLSSASASIISSLTLLMLPCIGFAMLMDLSGRFLLGTPIILIASVITL 595
 673 VNSLIDGLTAAALSTYSVIVYPCCFPMGFGPIP--NLICAEIPFTRVGLCIAICAF 730
 596 VLVNVLVGVTHVHAALSTISIVYVFCFFVM--GPIPGSSTSSVXIDIDRANAASIALAL 653
 731 TFWIGDIIVTYSLPVMLNAGIAGVFSIYAVVCLISVPVLPKVPETKMPLEVTTEPFA 790
 654 TFWIGDIIVTYSLPVMLNAGIAGVFSIYAVVCLISVPVLPKVPETKMPLEVTTEPFA 713
 791 VGAKQ 795
 714 VGAKQ 718

RESULT 9

Q96290 PRELIMINARY; PRT; 734 AA.
 AC Q96290;
 DT 01-FEB-1997 (Tremblrel. 02, Created)
 DT 01-FEB-1997 (Tremblrel. 02, last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, last annotation update)
 DE Sugar transporter.
 OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 RN
 RP SEQUENCE FROM N.A.
 RA Tjaden J., Neuhaus E.;
 RT "A new sugar transport protein from Arabidopsis thaliana";
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL; Z50752; CAA90628.1; -
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub transporter.
 DR InterPro; IPR001663; Sugar transp.
 DR InterPro; IPR005829; Sug_transporter.
 DR Pfam; PF00083; sugar_tr.1.
 DR PROSITE; PR00171; SUGTRNSPORT.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 DR Sugar transporter; Transmembrane; Transporter.
 SQ SEQUENCE 734 AA; 79462 MW; 1B009F5F457AF6E1 CRC64;
 Query Match 53.6%; Score 2183; DB 10; Length 734;
 Best Local Similarity 59.5%; Pred. No. 5,7e-140;
 Matches 449; Conservative 110; Mismatches 160; Indels 36; Gaps 14;
 QY 56 MSGAVLVAIVASIGNLQGMNATIAAVALYIKKEFOLONEPT-VEGLIVMSLIGATIV 114
 DB 1 MKGATLVLAATIGNPLQGMNATITAGAMVYINKLNL-PTVQGLVAMSLIGATVI 57
 QY 115 TTFSGLSDSISGRPMILISLIFPSGLIMSPNVVLLARVDFSGIANTVPL 174
 DB 58 TTCSGLSDSISGRPMILISLIFPSGLIMSPNVVLLARVDFSGIANTVPL 117
 QY 175 YISEIAPSEIRGLNTLTPQFSGSGMFLSYCMVFGMSLSPSPDMRIMGVLAIPSLFFPG 234
 DB 118 YISEIAPSEIRGLNTLTPQFSGSGMFLSYCMVFGMSLSPSPDMRIMGVLAIPSLFFPG 177
 QY 235 LTIPLPSFPMVLVSKGMAEAKVYLQKRGKDVSGELSLLEGLVEGDTISIEYIIG 294
 DB 178 LTIPLPSFPMVLVSKGMAEAKVYLQKRGKDVSGELSLLEGLVEGDTISIEYIIG 237
 QY 295 PATKADLVLDGDKK-QITLYGREGOSWILARPSKPIMLGSLVSLASRHSMSNOGVP 353
 DB 238 LEDHGDPTLTVEDDQIRLYGTENOSYLARPVPEQ--NSLGLRSRHGSLALNOGMI 294
 QY 354 LMDPIVTLFGSVHEMPOAGSGMRSTLFPNGSMFSTVDQ--HAKNEQMD--EENLRD 408
 DB 295 LMDPIVTLFGSVHEMPOAGSGMRSTLFPNGSMFSTVDQ--HAKNEQMD--EENLRD 354
 QY 409 DEEYASDAGGDYED--NLHSPILSRQATGAEGKDIYHGHGSGALSRRQTLLEGSD 465
 DB 355 NDDVATDDGAGDDDDSDNDLRSLMRSQRTSMQ-KDMIPHTSGTSLSRHSTLMQ-GN 412
 QY 466 GVSSTDDIGGQWOLAMKSEKGEENKRCGKFRVYLHOGVGVSGRGSIVSLPGGDIYFE 525
 DB 413 GVSSTDDIGGQWOLAMKSEKGEENKRCGKFRVYLHOGVGVSGRGSIVSLPGGDIYFE 525
 QY 526 GSEFHAALVYSQSLAFSGKLAEPHMSDAAMVHPSVAAKGRWKDLEPGVRALLVGV 585
 DB 461 GSEFHAALVYSQSLAFSGKLAEPHMSDAAMVHPSVAAKGRWKDLEPGVRALLVGV 515
 QY 586 GIQIQQPAGINGVLYTPQIIEAGVAVILSKFGLSSASASILISLTLLMLPCIGRA 645
 DB 516 GIQIQQPAGINGVLYTPQIIEAGVAVILSKFGLSSASASILISLTLLMLPCIGRA 575
 QY 646 MLMLNSGRRLLTGTPILLISLVILVSNLIDGTALHALLTSVSYLVYFCVMPG 705
 DB 576 MLMLNSGRRLLTGTPILLISLVILVSNLIDGTALHALLTSVSYLVYFCVMPG 635
 QY 706 PIPNLCABIFPTRVGLCAICAFPMWIGIIVYISLVPMALNIGLGVFSIYAVVGLI 765

DB 636 PFOSSVLKSSQADRGICIAICAMVFMWIGIIVYISLVPLSLISLHGVFSIYAAVCVI 695
 QY 766 SFVVFVFLAKVETKMPLEVFTEFFAVGAKQAAAA 800
 DB 696 SWIFVYMKVPETKMPLVITDYFAPGA-QAQA 729
 RESULT 10
 ID 08H87 PRELIMINARY; PRT; 714 AA.
 AC 08H87;
 DT 01-MAR-2003 (TREMblrel. 23, Created)
 DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)
 DE Putative sugar transporter protein.
 GN OUY52BD07.4.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; liliopsida; Poales; Poaceae;
 OC Euphorbiaceae; Oryzae; Oryza.
 NCBI_TaxID=39947;
 RN
 RP SEQUENCE FROM N.A.
 RA Wing R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Saeki C.,
 RA Currie J., Collura K.;
 RT "Rice Genomic Sequence";
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC099739; AAN17390.1; -
 DR Sugar transporter.
 SQ SEQUENCE 714 AA; 77883 MW; E6DA38A692106885 CRC64;
 Query Match 43.3%; Score 1763; DB 10; Length 714;
 Best Local Similarity 47.5%; Pred. No. 2.1e-111;
 Matches 377; Conservative 125; Mismatches 193; Indels 98; Gaps 14;
 QY 4 GSWLAVQPTFPDREERLLPSVVALPGPLPASCSSGSEPTSDILLEDKMGAVVA 63
 DB 9 GAW-----KATKQOVSPPSPISLSSVSESSLELSTTAD--GSTANAVLAA 54
 QY 64 IVASIGNLQGMNATIAAVALYIKKEFOLONEPTVEGLIVMSLIGATIVTTFSGLPLD 123
 DB 55 IVASIGNLQGMNATIAAVALYIKKEFOLONEPTVEGLIVMSLIGATIVTTFSGLPLD 114
 QY 124 SIGRPMILISLIFPSGLIMLSPNVVLLARVDFSGIANTVPLIYIETAPSE 183
 DB 115 SIGRPMILISLIFPSGLIMLSPNVVLLARVDFSGIANTVPLIYIETAPSE 174
 QY 184 IRGLNTLTPQFSGSGMFLSYCMVFGMSLSPSPDMRIMGVLAIPSLFFGLTI FYLPES 243
 DB 175 IRGLNTLTPQFSGSGMFLSYCMVFGMSLSPSPDMRIMGVLAIPSLFFGLTI FYLPES 234
 QY 244 PRWLVSCKRMAEAKVYLQKRGKDVSGELSLLEGLVEGDTISIEYIIGPATEADDL 303
 DB 235 PRWLVSCKRMAEAKVYLQKRGKDVSGELSLLEGLVEGDTISIEYIIGPATEADDL 293
 QY 304 VTDDGKEQITLYGREGOSWILARPSKPIMLGSLVSLASRHSMSNOGVP 363
 DB 294 QISINTETTKLYGHEGVTWFAFPFKGNV-----VESDHSPIPN--LLDPIVTLFD 343
 QY 364 SVHEMPOAGSGMRSTLFPNGSMFSTVDQAKAKNEQMDDEENLRDDEEVASDAGGDYED 423
 DB 344 SVHEMPOAGSGMRSTLFPNGSMFSTVDQAKAKNEQMDDEENLRDDEEVASDAGGDYED 417
 QY 424 NLHSPILSRQATGAEGKDIYHGHGSGALSRRQTLLEGSDVSTDIGGQWOLAMKWS 483
 DB 385 NLHSPILSRQATGAEGKDIYHGHGSGALSRRQTLLEGSDVSTDIGGQWOLAMKWS 417
 QY 484 EKGEENKRCGKFRVYLHOGVGVSGRGSIVSLPGGDIYFESEFHAALVYSQSLAFS 543
 DB 418 EKGEENKRCGKFRVYLHOGVGVSGRGSIVSLPGGDIYFESEFHAALVYSQSLAFS 458
 QY 544 KGLAEPHMSDAAMVHPSVAAKGRWKDLEPGVRALLVGVGIQIQQPAGINGVLY 601

```

Db      459 HSIG-----PAMVHPSKFNLTSTEGQSWEDLLQPGVKGLIVGTIQLIQLAGISGLY 512
Qy      602 YTPQLLEQAGAVVILSKFGLSSASASIISSLTLLMPCIFPAMLMDSRRFLIGT 661
Db      513 YTPQLLEQAGAGVILSKFGLSSASASIISSLTLLMPCIFPAMLMDSRRFLIGT 572
Qy      662 IPIILASVILVNSLIDIGTLAHLASTVSVIYVFCFVWGFGPIPNILCAIEFPTVR 721
Db      573 IPIILASVILVNSLIDIGTLAHLASTVSVIYVFCFVWGFGPIPNILCAIEFPTVR 632
Qy      722 GLCIAICAFTEWIGDIIVTYSLPVNLAIAGLVFSIYAVVCLISFVFLKVPETKMP 781
Db      633 NRCMSICTLFEWIVSIIYVAFPMVLSIGLIGVCGIYAVVCISFIFVLIKVPEKMP 692
Qy      782 LEVITEFPAVAK 794
Db      693 LAVIANSILAVGAR 705

RESULT 11
Q9SNK7 PRELIMINARY; PRT; 723 AA.
AC Q9SNK7;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Similar to sugar transporter protein.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrharioideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 3, PAC
   clone: p0043E01."
RL Submitted (Oct-1999) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AP000615; BAAB5398.1; -.
DR Gramene; Q9SNK7; -.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transp.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar_tr.1.
DR PRINTS; PR00171; SUGTRANSPORT.
DR PROSITE; PS00850; MFS.1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
DR KEGG; SUGAR_TRANSPORT; Transport.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 723 AA; 79061 MW; ECIE924FC4206BD8 CRC64;

Query Match 43.3%; Score 1763; DB 10; Length 723;
Best Local Similarity 47.5%; Pred. No. 2.1e-111;
Matches 377; Conservative 125; Mismatches 193; Indels 98; Gaps 14;

```

```

Db      184 MRGLNTLPQNGSIGMLSTYIMFVLSLTINPMRIMLGSLISFVFLIKVPEKMP 243
Qy      244 PRMIVSKRMABAKVLOKRGKODVSGELSLLEGLVGGDTSEETITGPATAADL 303
Db      244 PVLIVSKKRIEAKVNRMLRGNTNESEIAPFLQGLVDDDNVIEDMICHNDDEPD 302
Qy      304 VTDDGKEQITLYGPEEGSWIARPSKPIMGSGVLSIARSQSMVNSQVPMIDPIVTLFG 363
Db      303 QSIINTEITLYGHEBGTWFAKPPKGNV-----VESDHSPIPN---LDDPIVTLFD 352
Qy      364 SVHENMPOAGSMSTLFPNGSMFSVTDOAKNEQWDEENLHDDDEEYASDAGGDYED 423
Db      353 SIHGNI-----LMTPEFTSGMNSNDIEQ-PKTDLESQEDLDT-----DYED 393
Qy      424 NLHSPLSRQATGABGDIVHGRGSLNSRQRTLLGEGDDYSDTDIGGKQOLAMKWS 483
Db      394 DLGHLPLF-----HQSVM-----EGIDDACVNGMHIAMKEV 426
Qy      484 EKEGNGRKEGFKRVYLHQGVPGSRGSIYSLPGGDVFESEFVMAALVQSALFS 543
Db      427 QRENEFGOTODDFQOIFLQ-----GDIHQAGRVSHATLVSTPS-FH 467
Qy      544 KGLAEPMSDAAMVHPSE--VAAKGRWKDLFEGVRRALLVGVGIQTLQGFAGINGVLY 601
Db      468 HSIG-----PAMVHPSKFNLTSTEGQSWEDLLQPGVKGLIVGTIQLIQLAGISGLY 521
Qy      602 YTPQLLEQAGAVVILSKFGLSSASASIISSLTLLMPCIFPAMLMDSRRFLIGT 661
Db      522 YTPQLLEQAGAGVILSKFGLSSASASIISSLTLLMPCIFPAMLMDSRRFLIGT 581
Qy      662 IPIILASVILVNSLIDIGTLAHLASTVSVIYVFCFVWGFGPIPNILCAIEFPTVR 721
Db      582 IPIILASVILVNSLIDIGTLAHLASTVSVIYVFCFVWGFGPIPNILCAIEFPTVR 641
Qy      722 GLCIAICAFTEWIGDIIVTYSLPVNLAIAGLVFSIYAVVCLISFVFLKVPETKMP 781
Db      642 NRCMSICTLFEWIVSIIYVAFPMVLSIGLIGVCGIYAVVCISFIFVLIKVPEKMP 701
Qy      782 LEVITEFPAVAK 794
Db      702 LAVIANSILAVGAR 714

RESULT 12
Q8LJX7 PRELIMINARY; PRT; 746 AA.
AC Q8LJX7;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Putative sugar transporter.
GN SB234M12.4.
OS Sorghum bicolor (Sorghum) (Sorghum vulgare).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Sorghum.
OX NCBI_TaxID=4558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Btx623;
RA Song R., Liaca V., Messing J.;
RT "Mosaic Organization of Orthologous Sequences in Grass Genomes.";
RL Submitted (Jul-2002) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AF527809; AAM94321.1; -.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transp.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar_tr.2.
DR PRINTS; PR00171; SUGTRANSPORT.

```


RESULT 14

09XEV1 PRELIMINARY; PRT; 664 AA.
AC 09XEV1;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Putative sugar transporter protein.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrharioideae; Oryzaceae; Oryza.
OX NCBI_Taxid=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Liaca V., Lou A., Young S., Messing J.;
RT "Microarray in a gene-dense region in maize, sorghum, and rice";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; AF119222; AAD27676.1; -.
DR Gramene; O9XEV1; -.
DR InterPro; IPR005828; Sub transporter.
DR Pfam; PF00083; sugar tr; 1.
KW Sugar transport; Transmembrane.
SQ SEQUENCE 664 AA; 71851 MW; BA0CAB40C11C29E7 CRC64;

Query Match 15.3%; Score 621.5; DH 10; Length 664;
Best Local Similarity 27.5%; Pred. No. 8.9e-34;
Matches 193; Conservative 87; Mismatches 183; Indels 239; Gaps 21;

1 IRSSGMLAVQPFPT-----PDDBREBRLPSVVALPGLPPASCS----- 41
123 VRTPEQLCSSTSTHQSASNOQLPSPQSHRLRLHPRAI---PLPALAARSKLSP 178
42 -----SOBVTSSDD-----ILEDKMSGAVLVAIV 65
179 RYRGTGAPFRFRGAGARIIPSVAPDRPCHGESSVYLERALVDEMKSTYFSNAV 238
66 ASIGNLLOGMNAT-----AAAVLYIKKEPOLONEPTVEGLIVMSLIGATVTTGSP 120
239 VSIGTLLGMPFTVLGNPSAEANIMHMKKEFLANGCPSIDGIILAVSGSIAITVFGS 298
121 LSDSGRRPMLSSILVFFSGLIMMSNVVLLARVDFGIGLAVTVPYISER 180
299 LLDWLGRR-----
181 PSEIRGLINTLFPQSGGMPFLYCMVFQMSLSPSFDWIMLVLAIPSLFFGLTIFYL 240
307 -----IVSYCLIFMTLIPSENMKIMGALFAPSLVYFALLVPTL 347
241 PESPMVLVSKGMAAKVYLQRLRGKDV-----GELSLILEGLEVGDTSTIEYIIGPA 296
348 PESPMVLVSDGKISARISLQWLRGKDVDSAGEIALIADGMNMITETAVGAGVAV 407
297 TEAADVLTVDGKEQTLTYGPEEGSWIARPKGPIMGSVSLASRHSMNQSVPLMD 356
408 -----RSQSF-----LSTSTNQMSRSTFYWH-----LSD 432
357 PIVTLPFGSVHENPQAGSGMRSTLFPNFGMPSTVTOHAKNQMBEENLHRDEEYASDG 416
433 PLVDLGSIHESMSLQAG-RNSYFPVFN-FTNYEQBTSQGRNDLSQGRKAVSAB- 489
417 AGGDYEDNLHSPLSRQATGAEKDI-VHGHRSGLSMNRQTLGEGGCVSTDIIGG 475
490 -EGNNGDNLQASLIS-QVASAETNDINTSEGTSSSYLRH---GTSTGSLAQDLISS- 543
416 MQLAMKSEKGEENRKEGFRVYVLLHQGVGSGRSGIVSLPGGDVPEGEFPHAAAL 535
544 -----LHGH-----DIEEDEEELHIAL 561
536 VSQSLPFGSLAEPHMSDAAMVHPSEVAAKSGRMKDLPEPVRALVGVGIQLQFAG 595

DB 562 SSQRL-CAGHPRPQ---VRLSETADIKKRVLLQPEVRALTCGMILQIQ--- 614

QY 596 INGVLYTPQILBOAGAVILSKFGLSASASIIISLTLL 637
DB 615 -----TFNILLPDGAKYHWAQWQMDLCIRVILGSRVLL 650

RESULT 15

P66742 PRELIMINARY; PRT; 457 AA.

AC P66742;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Similar to metabolite transport proteins.
GN YMTG.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_Taxid=1423;
RN [1]
RP SEQUENCE FROM N.A.
RA LeLong C., Glaser P., Presecan E., Danchin A.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=96044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Berteiro M.G., Bessieres P., Bolojin A., Borchert S.,
RA Borries R., Boursier L., Brans A., Braun M., Brigelli S., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani U.J., Comerton I.F., Cummings N.J., Daniel R.A.,
RA Denzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entlan K.D., Ertington J., Fabre C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holseppel S., Hosono S., Hullo M.F., Itaya M., Jones J.,
RA Joris B., Karamata D., Kaashara Y., Klaert-Blanchard M., Klein C.,
RA Kobayashi Y., Koelter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapides A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigne C.,
RA Medina N., Meliado R.P., Mizuno M., Moesti D., Nakai S., Noack M.,
RA Noote D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Paresan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha B., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Serró S.J., Serró P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takayachi M., Tamakoshi A., Tanaka T., Terpiltra P., Tognoni A.,
RA Toesato V., Uchiyama S., Vandenbol M., Vannier F., Vassaretti A.,
RA Viari A., Wambut R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wiput A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the gram-positive bacterium Bacillus
subtilis".
RL Nature 390:249-256 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; Z92954; CAB07473.1; -.
DR EMBL; Z99122; CAB15600.1; -.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR007019; Prot_kinase.
DR InterPro; IPR005828; Sub transporter.
DR InterPro; IPR003663; Sugar_transpct.

DR InterPro: IPRO05829; Sug_transporter.
DR Pfam: PF00083; sugar_tr_1.
DR PRINTS: PRO0171; SUGTRANSFPORT.
DR TIGRFams: TIGR00879; Sp_1.
DR PROSITE: PS50850; MFS; 1.
DR PROSITE: PS00167; PROTEIN KINASE ATP; 1.
DR PROSITE: PS00216; SUGAR TRANSPORT_1; 1.
DR PROSITE: PS00217; SUGAR TRANSPORT_2; 1.
KW Transmembrane; Complete proteome
SQ SEQUENCE 457 AA, 49192 MW, E668EE0C4C8BEF7 CRC64

Query Match	14.4%	Score	587.5	DB	16	Length	457
Best Local Similarity	22.8%	Pred.	No.	1.1e-31			
Matches	168	Conservative	102	Mismatches	162	Indels	305
						Gaps	12

[illegible]

Search completed: January 5, 2004, 18:57:50
Job time : 105 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 15:48:00 ; Search time 74 Seconds

(without alignments)
1715.963 Million cell updates/sec

Title: US-10-051-909-32

Perfect score: 4075
Sequence: 1 IRSGSWLAVQPFPTDLDR.....PLEVTEFFAVGAKQAAKA 800

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq.19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3517	86.3	747	23	AAU97201
2	3517	86.3	747	24	ABU08326
3	2674	65.6	737	23	AAU97204
4	2674	65.6	737	24	ABU08329
5	2584	63.4	729	23	ABR93352
6	2275	55.8	734	23	ABR91078
7	2245	55.1	729	23	ABR92622
8	1670.5	41.0	486	23	AAU97205
9	1670.5	41.0	486	24	ABU08330

10	1424	34.9	345	23	AAU97207
11	1424	34.9	345	24	ABU08332
12	943	23.1	228	23	AAU97208
13	943	23.1	228	24	ABU08333
14	562	13.8	131	23	AAU97203
15	562	13.8	131	24	ABU08328
16	554.5	13.6	487	24	ABG73334
17	548	13.4	488	24	ABP98504
18	520.5	12.8	555	21	AAQ9529
19	520.5	12.8	582	21	AAQ9528
20	496	12.2	523	23	AAU97212
21	496	12.2	523	24	ABU08337
22	493.5	12.1	491	21	AAQ32072
23	493.5	12.1	508	21	AAQ32071
24	489.5	12.0	513	23	AAU97210
25	489.5	12.0	513	24	ABU08335
26	489	12.0	529	23	AAU97214
27	489	12.0	529	24	ABU08339
28	485	11.9	510	23	AAU97211
29	485	11.9	510	24	ABU08336
30	481	11.8	513	23	ABU09681
31	468	11.5	551	20	AAV50799
32	467	11.5	478	21	AAQ15416
33	467	11.5	493	21	AAQ15415
34	467	11.5	546	21	AAQ15414
35	463.5	11.4	539	23	AAU97213
36	463.5	11.4	539	24	ABU08338
37	457.5	11.2	466	21	AAQ32073
38	453.5	11.1	480	22	AAQ00107
39	451	11.1	464	23	ABP52163
40	448	11.0	474	19	AAW58862
41	444	10.9	491	22	AAQ99949
42	443	10.9	508	21	AAQ29530
43	442	10.8	491	21	AAQ12594
44	429.5	10.5	472	23	ABP52164
45	425	10.4	629	22	AAQ66935

ALIGNMENTS

RESULT 1	AAU97201	standard; Protein; 747 AA.
ID	AAU97201	
AC	AAU97201;	
XX	27-AUG-2002	(first entry)
DT		
XX		
DE	Corn Arabidopsis thaliana-like sugar transport protein.	
KW	Corn; Arabidopsis thaliana-like sugar transport protein;	
KW	carbohydrate transport; grain filling; annual field crop;	
KW	plant.	
XX		
OS	Zea mays.	
XX		
FH	Key	Location/Qualifiers
FT	Misc-difference	129
FT	/label=	Unknown
FT	Misc-difference	133..134
FT	/label=	Unknown
FT	Misc-difference	144
FT	/label=	Unknown
FT	Misc-difference	178
FT	/label=	Unknown
FT	Misc-difference	207
FT	/label=	Unknown
FT	Misc-difference	218
FT	/label=	Unknown
FT	Misc-difference	220
FT	/label=	Unknown
FT	Misc-difference	236

Portion of a wheat
Wheat sugar transp
Portion of a wheat
Wheat sugar transp
Portion of a rice
Rice sugar transp
Consensus amino ac
PPAM consensus seq
Arabidopsis thalia
Arabidopsis thalia
Soybean Beta vulga
Soybean sugar tran
Arabidopsis thalia
Arabidopsis thalia
Corn Beta vulgaris
Corn sugar transpo
Wheat sugar transp
Wheat sugar transp
Rice Beta vulgaris
Rice sugar transp
Amino acid sequenc
Spinach glucose tr
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Wheat sugar transp
Wheat sugar transp
Arabidopsis thalia
Sugar transporter
E. coli galactose-
T. halophilus xyul
C. glutamicum prote
Arabidopsis thalia
Brevibacterium lac
E. coli arabinose-
Human GLUTX. Hom

/label= unknown

Applicant

FT US6383776-B1.
 XX 07-MAY-2002.
 XX 14-APR-1999; 99US-0291922.
 XX 24-APR-1998; 98US-083044P.
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX Allen SM, Hiltz WD, Kinney AJ, Tingey SV;
 XX WPI; 2002-451386/48.
 DR N-PSDB; ABKS1962.
 XX New nucleic acid encoding plant sugar-transport proteins, useful for
 PT preparing transgenic plants with altered carbohydrate distribution -
 XX
 PS Claim 1; Fig 1; 54pp; English.
 CC The present invention relates to the isolation of plant polynucleotide
 CC sequences encoding an Arabidopsis thaliana-like sugar transport
 CC protein or Beta vulgaris-like sugar transport protein. The
 CC polynucleotide sequences are useful for altering the level of sugar
 CC transport proteins in plants, i.e. for control of carbohydrate transport
 CC and distribution in plant cells, e.g. during grain filling of annual
 CC field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
 CC carbohydrate flows and sugar transport. The polynucleotide
 CC sequences can also be used to isolate cDNA sequences and genes that
 CC encode homologues of the new proteins. The present sequence
 CC represents a corn Arabidopsis thaliana-like sugar transport protein.
 SQ Sequence 747 AA;

Query Match Best Local Similarity 91.6%; Score 3517; DB 23; Length 747;

Matches 684; Conservative 27; Mismatches 34; Indels 2; Gaps 2;

QY 56 MSGATLVAVASISGILQGMNDNATIAAALVYIKKEFOLNEPTVEGLIYMSLIGATVIT 115
 DB 1 MCGATMVAVASISGILQGMNDNATIAAALVYIKKEFOLNEPTVEGLIYMSLIGATVIT 60
 QY 116 TFSGLSDSISGRPMILISSILYFPGSLIMSPVYVLLARFVDFGICGLAVTLVPLV 175
 DB 61 TSPGRADCVGRPMILISSILYFPGSLIMSPVYVLLARFVDFGICGLAVTLVPLV 120
 QY 176 ISEIAPSEIR-GLMTLTPQFSG-SGGMPLSYCMVFCMSISPSPPRIMGVATISLPPF 233
 DB 121 ISEIAPSEIR-GLMTLTPQFSG-SGGMPLSYCMVFCMSISPSPPRIMGVATISLPPF 180
 QY 234 GLTTFYLPESPRMILVSKGRMAEAKVLOKLRGKDVSGELSLILBGLLEVGGDTISIEYI 293
 DB 181 GLTTFYLPESPRMILVSKGRMAEAKVLOKLRGKDVSGELSLILBGLLEVGGDTISIEYI 240
 QY 294 GPATPAADLVTDGKEQITLYGPEGGOSWARPSPKIPMIGSVISLASRHGSMNQSV 353
 DB 241 GPATPAADLVTDGKEQITLYGPEGGOSWARPSPKIPMIGSVISLASRHGSMNQSV 300
 QY 354 LMDPIVTLFGSVHNMPPQAGSMRSTLPNFGSMFESVTDQAKNOMBEENIHRDEEYA 413
 DB 301 LMDPIVTLFGSVHNMPPQAGSMRSTLPNFGSMFESVTDQAKNOMBEENIHRDEEYA 360
 QY 414 SDGAGGEDVNDLHSPILSRQATGABKDIIVHGHGHSALSMRQTLLEGGGVSTDIG 473
 DB 361 SDGAGGEDVNDLHSPILSRQATGABKDIIVHGHGHSALSMRQTLLEGGGVSTDIG 420
 QY 474 GGMOLAMKSEKSEKNGRKGKPKVYVYHQBGVPCGRGSIYSLPGGGDFVFGSEFVAA 533
 DB 421 GGMOLAMKSEKSEKNGRKGKPKVYVYHQBGVPCGRGSIYSLPGGGDFVFGSEFVAA 480
 QY 534 ALVSGSALFSGKGLAPRMSDAMVHPSEVAAGSKWKLDFEGVRRALLVGVGIOLQOF 593

DB 481 ALVSGSALFSGKGLAPRMSDAMVHPSEVAAGSKWKLDFEGVRRALLVGVGIOLQOF 540
 QY 594 AGINGVLYTPQILLEGAGVAVILSKFGLSSASASILISLTLTLMPCIGFAMLMDSG 653
 DB 541 AGINGVLYTPQILLEGAGVAVILSKFGLSSASASILISLTLTLMPCIGFAMLMDSG 600
 QY 654 RRFLLGTIPILIASVILVNSNIDGTLAHLSTVSIVVRCFCFMRGPIPNILCA 713
 DB 601 RRFLLGTIPILIASVILVNSNIDGTLAHLSTVSIVVRCFCFMRGPIPNILCA 660
 QY 714 EIPFRVAGLCAICAFPMIGDIIVTYSLEPMLNATIGAGVFIYAVCLISFVFLK 773
 DB 661 EIPFRVAGLCAICAFPMIGDIIVTYSLEPMLNATIGAGVFIYAVCLISFVFLK 720
 QY 774 VPETKGMPLLEVITEFFPAVGAQAAKA 800
 DB 721 VPETKGMPLLEVITEFFPAVGAQAAKA 747

RESULT 2
 ABU08326 standard; Protein; 747 AA.
 AC ABU08326;
 DT 29-MAY-2003 (first entry)
 XX Corn sugar transport protein #1.
 XX Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
 KM plant sugar transport protein; carbohydrate transport; soybean;
 KM carbohydrate distribution; plant.
 XX Zea mays.
 OS US2002178468-A1.
 XX
 PN 28-NOV-2002.
 XX
 PD 17-JAN-2002; 2002US-0051902.
 XX
 PF 24-APR-1998; 98US-083044P.
 XX
 PR 14-APR-1999; 99US-0291922.
 XX

Applicant

PA (ALLEE) ALLEN S M.
 PA (HITZ) HITZ W D.
 PA (KINNEY) KINNEY A J.
 PA (TINGEY) TINGEY S V.
 PI Allen SM, Hiltz WD, Kinney AJ, Tingey SV;
 XX
 XX WPI; 2003-340957/32.
 DR N-PSDB; ABX93198.
 XX
 PT Novel plant sugar transport proteins and nucleic acid encoding the
 PT protein useful for producing transgenic plants having altered levels of
 PT sugar transport protein -
 XX
 PS Claim 5; Fig 1; 56pp; English.
 CC The present invention relates to the isolation of Arabidopsis
 CC thaliana-like or Beta vulgaris-like sugar transport proteins, and
 CC the polynucleotide sequences encoding them. The plant sugar transport
 CC proteins of the invention have been isolated from corn, rice, soybean,
 CC and wheat. The polypeptides of the invention may be used for altering
 CC the level of expression of a sugar transport protein in a host cell,
 CC by transforming a host cell with a chimeric construct encoding all,
 CC or a portion of the sugar transport protein, in sense or antisense
 CC orientation. Particularly, the polypeptides may provide a means to
 CC control carbohydrate transport and distribution in plants.
 CC ABU08326-ABU08333 represent Arabidopsis thaliana-like sugar
 CC transport proteins.

XX	Sequence	747 AA:	
SQL	Query Match	86.3%; Score 3517; DB 24; Length 747;	
	Best Local Similarity	91.6%; Pred. No. 0;	
	Matches	684; Conservative 27; Mismatches 34; Indels 2; Gaps 2;	
QY	56	MSGAVLVAVASISGLTLQGMNATIAAVALYIKKEFOLENEPTVEGLIVMSLIGATVIT	115
DB	1	MCGAVMAVIAISIGTLQGMNATIAAGAVLYIKKEFNQSEPLIGLIVAMFLITATVIT	60
QY	116	TFSGPLSDSISGRPMILISILYFSGILMSPNVYLLARFVDFGIGLATVLPVLY	175
DB	61	TSPGRADCVGRPMVLASAVLYFVSGVLMWAPVYIILLARLIDGFIGLATVLPVLY	120
QY	176	ISELAPSEIR-GLNTLTPQFSG-SCGMFLSYCMVGMISLSPEDRMIMGLVLAISLFF	233
DB	121	ISELAPHXSMGXNNTLPQFIGVXGMFLSYCMVGMISLSPEDRMIMGLVLAISLXYF	180
QY	224	GLTFYLPESPRLVSKGRMAEAKVYLOKRGKDVSGELSLLEGLVGGDTISEYII	293
DB	181	GLTFYLPESPRLVSKGRMAEAKVYLOKRGKEDVSKEXALLVGLVGDTRIXEYII	240
QY	294	GPATEAADLVYTDGKEQITLYPREGQSWIARPSKGPIMLSVLSASRHSVMNQSP	353
DB	241	GPATEAADLVYTDGKEQITLYPREGQSWIARPSKGPIMLSVLSASRHSVMNQSP	300
QY	354	LMDPVLVPGSVHEHMPQAGSGMSTLPNFGSMFSTVDQAKNQMBENLHRDEEYA	413
DB	301	LMDPVLVPGSVHEHMPQAGSGMSTLPNFGSMFSTVDQAKNQMBENLHRDEEYA	360
QY	414	SDGAGDYEDNLHSPILSRQATGAEKQIVHGHGSAALSMRQTLLEGGDVSSTIG	473
DB	361	SDGAGDYEDNLHSPILSRQATGAEKQIVHGHGSAALSMRQTLLEGGDVSSTIG	420
QY	474	GGMQLAMKSEKENGKRGKGFVYVYHQBEGVPSGRGSIYSLPGGDFEGSEFVHA	533
DB	421	GGMQLAMKSEKENGKRGKGFVYVYHQBEGVPSGRGSIYSLPGGDFEGSEFVHA	480
QY	534	ALVSGSALFSGKLAEPMSDAAVHPSEVAAKSGRWKOLFEGVRRALLVVGIOIILQOF	593
DB	481	ALVSGSALFSGKLAEPMSDAAVHPSEVAAKSGRWKOLFEGVRRALLVVGIOIILQOF	540
QY	594	AGINCVLYTTPQIIEQAGVAVILSKFGLSSASASTLISLTLLMPCIGFAMLLMDISG	653
DB	541	AGINCVLYTTPQIIEQAGVAVILSKFGLSSASASTLISLTLLMPCIGFAMLLMDISG	600
QY	654	RRFLILGTIPILASLVILVSNLIDGLTAHALSTVSIVYFCCFVWGFGPIINILCA	713
DB	601	RRFLILGTIPILASLVILVSNLIDGLTAHALSTVSIVYFCCFVWGFGPIINILCA	660
QY	714	EIPFRVAGLCIAICFTFWIGDIIVYVSLPVLNAIGLAVSIVAVVCLISFVFLK	773
DB	661	EIPFRVAGLCIAICFTFWIGDIIVYVSLPVLNAIGLAVSIVAVVCLISFVFLK	720
QY	774	VPETGMPLVETTERFAVCAKQAAKA 800	
DB	721	VPETGMPLVETTERFAVCAKQAAKA 747	

XX	OS	Glycine max.	
XX	PN	US683776-B1.	
XX	PD	07-MAY-2002.	
XX	PF	14-APR-1999; 99US-0291922.	
XX	PR	24-APR-1998; 98US-083044P.	
XX	PA	(DUPO) DU PONT DE NEMOURS & CO B I.	
XX	PI	Allen SM, Hiltz WD, Kinney AJ, Tingey SV;	
XX	DR	WPI; 2002-451386/48.	
XX	DR	N-PSDB; ABK51965.	
XX	PT	New nucleic acid encoding plant sugar-transport proteins, useful for	
XX	PT	preparing transgenic plants with altered carbohydrate distribution -	
XX	PS	Claim 1; Fig 1; 54pp; English.	
XX	CC	The present invention relates to the isolation of plant polynucleotide	
XX	CC	sequences encoding an Arabidopsis thaliana-like sugar transport	
XX	CC	protein or beta vulgaris-like sugar transport protein. The	
XX	CC	polynucleotide sequences are useful for altering the level of sugar	
XX	CC	transport proteins in plants, i.e. for control of carbohydrate transport	
XX	CC	and distribution in plant cells, e.g. during grain filling of annual	
XX	CC	field crops (e.g. corn, rice, soybeans, and wheat), and, for studying	
XX	CC	carbohydrate flows and sugar transport. The polynucleotide	
XX	CC	sequences can also be used to isolate cDNA sequences and genes that	
XX	CC	encode homologues of the new proteins. The present sequence	
XX	CC	represents a soybean Arabidopsis thaliana-like sugar transport protein.	
XX	SQL	Sequence 737 AA:	
	Query Match	65.6%; Score 2674; DB 23; Length 737;	
	Best Local Similarity	69.9%; Pred. No. 276-249;	
	Matches	522; Conservative 92; Mismatches 115; Indels 18; Gaps 8;	
QY	56	MSGAVLVAVASISGLTLQGMNATIAAVALYIKKEFOLENEPTVEGLIVMSLIGATVIT	115
DB	1	MKGAVLVAVASISGLTLQGMNATIAAGVYIKKEDLALGT--THERLVVGMISLIGATVIT	58
QY	116	TFSGPLSDSISGRPMILISILYFSGILMSPNVYLLARFVDFGIGLATVLPVLY	175
DB	59	TCSGPIADWLARRRPMIITISVLYFGLVIMLSPNVYLLARLIDGFIGLATVLPVLY	118
QY	176	ISELAPSEIRGLNLTLPQFSGGGMFLSYCMVFGMSLSPSDRMIMGLVLAISLFFGL	235
DB	119	ISELAPSEIRGLNLTLPQFSGGGMFLSYCMVFGMSLSPSDRMIMGLVLAISLFFGL	178
QY	236	TIFPLPESPRMLVSKGRMAEAKVYLOKRGKDVSGELSLLEGLVGGDTISEEYIIGP	295
DB	179	TIFPLPESPRMLVSKGRMAEAKVYLOKRGKEDVSGEALVLBGIGIGDTISEEYIIGP	238
QY	296	ATEAADLVYTDGKEQITLYPREGQSWIARPSKGPIMLSVLSASRHSVMNQSPVLM	355
DB	239	ADDVADGHEHATEKDKIRLYSGOAGLSWLSKPVVGSSISG---LASHHGSITINQSMPLM	294
QY	356	DPVYTLFSGSVHEHMPQAG--GSMSTLPNFGSMFSTVDQAKNQMBENLHRDEEYA	413
DB	295	DPVYTLFSGSVHEHMPQAG--GSMSTLPNFGSMFSTVDQAKNQMBENLHRDEEYA	354
QY	414	SDGAGDYEDNLHSPILSRQATGAEKQIVHGHGSAAL-SMRQTLLEGG-GLGVSSSTD	471
DB	355	SDATRGDSDDLHSPILSRQATGAEKQIVHGHGSAAL-SMRQTLLEGG-GLGVSSSTD	413
QY	472	GGGQLAMKSEKENGKRGKGFVYVYHQBEGVPSGRGSIYSLPGGDFVFGSEFVH	531
DB	414	GGGQLAMKSEKENGKRGKGFVYVYHQBEGVPSGRGSIYSLPGGDFVFGSEFVH	466

RESULT 3
 AAU97204
 ID AAU97204 standard; Protein; 737 AA.
 AC AAU97204;
 AC 27-AUG-2002 (first entry)
 DT Soybean Arabidopsis thaliana-like sugar transport protein.
 XX Soybean; Arabidopsis thaliana-like sugar transport protein;
 XX carbohydrate transport; grain filling; annual field crop;
 KW plant.

QY 532 AALVVSQSALFESKGLAEPRMSDAAVHPSEVAAGSRWKDLFEPGVRRALLVGVGIQIQQ 591
 DB 467 AALVVSQPALYXSKLIDGHPVPAPVHPSERASKSPWALIEPGKHALVVGVIQIQQ 526
 QY 592 QPAGINGVLYYTPQILEQAGVAVILSKFGLSSASASILISLTLLMLPCTGFAMLLMDL 651
 DB 527 QPAGINGVLYYTPQILEQAGVAVILSDIGISESASFLISATTTLLMLPCTGFAMKMDV 586
 QY 652 SGRFRLLLGTPIILASVILVSNLIDGLTAHALSTVSIVYVFCFVWGFGPIPNIL 711
 DB 587 SGRFRLLLGTPIILASVILVSNLIDGLTAHALSTVSIVYVFCFVWGFGPIPNIL 646
 QY 712 CAEIFPTRVRGLCIAICAFTEFWIGIIVTYSLPVMLNAGLAVSIYAVVCLISFVFEV 771
 DB 647 CSEIFPTRVRGLCIAICAFTEFWIGIIVTYSLPVMLNAGLAVSIYAVVCLISFVFEV 706
 QY 772 LKVPETKGMPLFVITEFFAVGAKQA 798
 DB 707 LKVPETKGMPLFVITEFFAVGAKQA 733

RESULT 4
 ABU08329 standard; Protein; 737 AA.
 ID ABU08329 standard; Protein; 737 AA.
 AC ABU08329;
 XX
 DT 29-MAY-2003 (first entry)
 XX
 DE Soybean sugar transport protein #1.
 XX
 KW Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
 KW plant sugar transport protein; carbohydrate transport; soybean;
 KW carbohydrate distribution; plant.
 XX
 OS Glycine max;
 XX
 PN US2002178468-A1.
 XX
 PD 28-NOV-2002;
 XX
 PF 17-JAN-2002; 2002US-0051902.
 XX
 PR 24-APR-1998; 98US-083044P.
 PR 14-APR-1999; 99US-0291922.
 XX
 PA (ALLEN/) ALLEN S M.
 PA (HITZ/) HITZ W D.
 PA (KINNEY/) KINNEY A J.
 PA (TINGE/) TINGEY S V.
 XX
 PI Allen SM, Hiltz WD, Kinney AJ, Tingey SV;
 XX
 DR WPI; 2003-340957/32.
 DR N-PSDB; ABX93201.
 XX
 PT Novel plant sugar transport proteins and nucleic acid encoding the
 PT protein useful for producing transgenic plants having altered levels of
 PT sugar transport protein -
 XX
 PS Claim 5; Fig 1; 56pp; English.

CC The present invention relates to the isolation of Arabidopsis
 CC thaliana-like or Beta vulgaris-like sugar transport proteins, and
 CC the polynucleotide sequences encoding them. The plant sugar transport
 CC proteins of the invention have been isolated from corn, rice, soybean,
 CC and wheat. The polypeptides of the invention may be used for altering
 CC the level of expression of a sugar transport protein in a host cell,
 CC by transforming a host cell with a chimeric construct encoding all,
 CC or a portion of the sugar transport protein, in sense or antisense
 CC orientation. Particularly, the polypeptides may provide a means to
 CC control carbohydrate transport and distribution in plants.
 CC ABU08326-ABU08333 represent Arabidopsis thaliana-like sugar

CC transport proteins.
 XX
 SQ Sequence 737 AA;
 Query Match 65.6%; Score 2674; DB 24; Length 737;
 Best Local Similarity 69.9%; Pred. No. 2.7e-249;
 Matches 522; Conservative 92; Mismatches 115; Indels 18; Gaps 8;

QY 56 MSGAVLVAIVASISNLLQGMNATITAAVLYIKKEPOLQNEPTEVGLIVSMIGATITV 115
 DB 1 MKGAVLVAIVASISNLLQGMNATITAGANGYIKKDLALGT--TWERLVAWGSLIGATITV 58
 QY 116 TFSGPLSGISGRARRMMLSSILYFSGIIMMSNNVYLLARVDDGIGLAAVTPLY 175
 DB 59 TFSGPLADWIGRRRPMMLISSVLYFLGLVLMMSNNVYLLCARLLDGGIGLAAVTPLY 118
 QY 176 ISEIAPSEIRGLNLTLPQFSGSGMPLSYCMVFCMSLSPSPDMRIMLVLAIPSLFFPGL 235
 DB 119 ISEIAPSEIRGLNLTLPQFSGSGMPLSYCMVFCMSLSPSPDMRIMLVLAIPSLFFPGL 178
 QY 236 TIFPLPESPRMLVSKGRMAEAKVYLQKRGKDVSGELSLLEGLVVGDTSEIEYIIGP 295
 DB 179 TIFPLPESPRMLVSKGRMLEAKVYLQKRGREDVSGEMALLVEGLIGDTSIEYIIGP 238
 QY 296 ATEAADLVTDGDEKQITTLVPREGSGMIAPSKGPIMLGSLVSLARHSGMNSQVPLM 355
 DB 239 ADVDAADGHEHATEKDKRLTLYGSGQGLSLMSPKVTGQSSIG--LASHRSGIINQSMPLM 294
 QY 356 DPIVTLFSGVHENMPQAG--GSMRSTLFPNFGSMFVTDQHAKEQWDEENLHRDDEEYA 413
 DB 295 DPLVTLRGSIHKKLPETGARGSMRSTLFPNFGSMFVTDQHAKEQWDEENLHRDDEEYA 354
 QY 414 SDGAGGVEDNLHSPILSRQATGAEKDIYVHGRGSL-SMRQTLIGS-GDGVSTG 471
 DB 355 SDATRGSDSDNLHSPILSRQATGAEKDIYVHGRGSL-SMRQTLIGS-GDGVSTG 413
 QY 472 IGGGMOLAMKSEKEGNGRKEGPKRVYLHQEGVPSRSGISVLPQGGVFFSGSEPVH 531
 DB 414 IGGGMOLAMKWTDK-GBDGKQGGPKRVYLHEBVSASRRGSIYIPBG-----EFVQ 466
 QY 532 AALVVSQSALFESKGLAEPRMSDAAVHPSEVAAGSRWKDLFEPGVRRALLVGVGIQIQQ 591
 DB 467 AALVVSQPALYXSKLIDGHPVPAPVHPSERASKSPWALIEPGKHALVVGVIQIQQ 526
 QY 592 QPAGINGVLYYTPQILEQAGVAVILSKFGLSSASASILISLTLLMLPCTGFAMLLMDL 651
 DB 527 QPAGINGVLYYTPQILEQAGVAVILSDIGISESASFLISATTTLLMLPCTGFAMKMDV 586
 QY 652 SGRFRLLLGTPIILASVILVSNLIDGLTAHALSTVSIVYVFCFVWGFGPIPNIL 711
 DB 587 SGRFRLLLGTPIILASVILVSNLIDGLTAHALSTVSIVYVFCFVWGFGPIPNIL 646
 QY 712 CAEIFPTRVRGLCIAICAFTEFWIGIIVTYSLPVMLNAGLAVSIYAVVCLISFVFEV 771
 DB 647 CSEIFPTRVRGLCIAICAFTEFWIGIIVTYSLPVMLNAGLAVSIYAVVCLISFVFEV 706
 QY 772 LKVPETKGMPLFVITEFFAVGAKQA 798
 DB 707 LKVPETKGMPLFVITEFFAVGAKQA 733

RESULT 5
 ABB93252
 ID ABB93252 standard; Protein; 729 AA.
 AC ABB93252;
 XX
 DT 31-MAY-2002 (first entry)
 XX
 DE Herbicidally active polypeptide SEQ ID NO 2463.
 XX
 KW Herbicidal; plant; agriculture; herbicide.

06 Arabidopsis thaliana.
 XX
 PN WO200210210-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 28-AUG-2001; 2001WO-EP09892.
 XX
 PR 28-AUG-2001; 2001WO-EP09892.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Tietjen K, Weidler M;
 XX
 DR WPI; 2002-269010/31.
 XX
 PT Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms -
 XX
 PS Claim 5; SEQ ID NO 2463; 261pp + Sequence listing; English.
 XX
 CC The invention relates to identifying target proteins
 CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
 CC aligning and comparing nucleic acid or amino acid sequences from plant
 CC with nucleic acid or amino acid sequences from non-plant organisms using
 CC suitable search parameters, where plant sequences having an E-value
 CC greater by a factor of 3 than the E-value of most similar non-plant
 CC sequences are selected. The polypeptides or nucleic acids encoding them
 CC are useful for identifying modulators. The identified modulators are
 CC useful as herbicides.
 XX
 SQ Sequence 729 AA.

Query Match 63.4%; Score 2584; DB 23; Length 729;
 Best Local Similarity 68.8%; Pred. No. 1.4e-240;
 Matches 512; Conservative 89; Mismatches 123; Indels 20; Gaps 8;

QY 56 MSGAVLVAIVASIGNLLQGMNDATIAAVLYIKKEFOLNEPTVEGLIVSMIGATIVT 115
 DB 1 MSGAVLVAIVAAVGNLLQGMNDATIAAGAVLYIKKEFNEISNPSVEGLIVAMSLIGATLIT 60
 QY 116 TFSGLSDSISGRPMILISSIIFYPSGLIMLSPNVYLLARFDVREGISGLAVTLVPIY 175
 DB 61 TCSGGADWLGRRPMLISSIIFYPSGLIMLSPNVYLLARFDVREGISGLAVTLVPIY 120
 QY 176 ISEIAPSEIRGLINTLPQPSGSGMFLSYCMYFGMSISPSPMRIMLGTLAIPSLFFGL 235
 DB 121 ISETAPPEIRGLINTLPQFTSGGMLSYCMYFGMSISPSPMRIMLGTLAIPSLVFFPL 180
 QY 236 TIFYPESPFRMLVSKGRMAEAKVKVLQKRGKDVSGELSLLEGLVEGDTISIEYITGP 295
 DB 181 TIFYPESPFRMLVSKGRMAEAKVKVLQKRGKDVSGELSLLEGLVEGDTISIEYITGP 240
 QY 236 AFEADADLVDDDKQITLYGPEEGSWTARSKGIIMGLSVLSASRRGS-MVNVSVL 354
 DB 241 ADEVTDHDIADVAKQIKLYGABEGISWVARPVK---GSTMSVLSIRSGTMSRQGL 296
 QY 355 MDPIVTLFGSVHNNPQAGSNRSTLFPNFGSMFSTVDQAKNQEWDENLHRDDEEYAS 414
 DB 297 IDPLVTLRGSVHEKMDT-GSMRSALFPHFGSMFSGNQPHEDMDENLVGEGEDYPS 355
 QY 415 DGAGDVEDNLSPLISROATGAEGKDIVHGHRSALSMBRQTLLEGSGDGVSTDIG 474
 DB 356 D-HGDDSEDDLSPLISROATGAEGKDIVHGHRSALSMBRQTLLEGSGDGVSTDIG 413
 QY 475 GMDLAKMSKSGENKRGKGRFVYVHGEVPGSRGIVLPFGGCVDFEESSEFTHAA 534
 DB 414 GMDVAKMTKTERBESQKE-----EGPFGSRGIVSPGSDGTGE-ADFOQAA 462
 QY 535 LVSGSALFSGGLAEPFMSDAAMVHPSEVAKGSRWMDLPEFGRRALLVGVGIQILQCPA 594

DB 463 LVSGPALYSKDLKEHTIGPAMVHPS-TRKGIWHDLDHDPVKRALVVGVLQILQOPS 521
 QY 595 GINGVLYTTPQILBOAGAVILSKPGLSASASITILSLTLLMPCIGPAMLLMDLSGR 654
 DB 522 GINGVLYTTPQILBOAGAVILSKPGLSASASITILSLTLLMPCIGPAMLLMDLSGR 581
 QY 655 RFLIGTIPILIASLVILVSNLIDLTALHALSTVSIVYVFCFVWGFGFIPNITLAE 714
 DB 582 RFLIGTIPILIASLVILVSNLIDLTALHALSTVSIVYVFCFVWGFGFIPNITLAE 641
 QY 715 IFPFRVRLCIAICAFIFWIDIVTYSLPVMAAIGLAGVPSIYAVVCLISFVFLKV 774
 DB 642 IFPFRVRLCIAICAFIFWIDIVTYSLPVMAAIGLAGVPSIYAVVCLISFVFLKV 701
 QY 775 PETKMPLEVTPEFPAKQAAA 798
 DB 702 PETKMPLEVTPEFPAKQAAA 725

RESULT 6
 ABB91078
 ID ABB91078 standard; Protein; 734 AA.

XX ABB91078;
 AC ABB91078;
 XX
 DT 31-MAY-2002 (first entry)
 XX
 DE Herbicidally active polypeptide SEQ ID NO 289.
 XX
 KW Herbicidal; plant; agriculture; herbicide.
 XX
 OS Arabidopsis thaliana.
 XX
 PN WO200210210-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 28-AUG-2001; 2001WO-EP09892.
 XX
 PR 28-AUG-2001; 2001WO-EP09892.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Tietjen K, Weidler M;
 XX
 DR WPI; 2002-269010/31.
 XX
 PT Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms -
 XX
 PS Claim 5; SEQ ID NO 289; 261pp + Sequence listing; English.
 XX
 CC The invention relates to identifying target proteins
 CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
 CC aligning and comparing nucleic acid or amino acid sequences from plant
 CC with nucleic acid or amino acid sequences from non-plant organisms using
 CC suitable search parameters, where plant sequences having an E-value
 CC greater by a factor of 3 than the E-value of most similar non-plant
 CC sequences are selected. The polypeptides or nucleic acids encoding them
 CC are useful for identifying modulators. The identified modulators are
 CC useful as herbicides.
 XX
 SQ Sequence 734 AA;
 XX
 Query Match 55.8%; Score 2275; DB 23; Length 734;
 Best Local Similarity 61.2%; Pred. No. 1.2e-210;
 Matches 462; Conservative 111; Mismatches 146; Indels 36; Gaps 14;
 QY 56 MSGAVLVAIVASIGNLLQGMNDATIAAVLYIKKEFOLNEPT-VEGLIVSMIGATIV 114
 DB 1 MKGATLVALLATIGNFLQGMNDATIAAGAVIYNINDLN---PTSVQGLVAMSLIGATIV 57

QY	115	ITFSGPLSNISGRRLPLLSSITLYFSGGLIMSPNVYUILLARPDPGPGIGLAUVTYPL	174
Db	58	TTCSGSIISMLGRRLPLLSSVWYFPCGILIMSPNVYUILLARPDPGPGIGLAUVTYPL	117
QY	175	YISEIAPSEIRGLANTLPQFSGSGGWFSLCYWFGNLSPPSPMRIMLGYLAPSLFFFG	234
Db	118	YISEIAPSEIRGLANTLPQFSGSGGWFSLCYWFTMSLSNPSBMRANLGVLSPLLYLF	177
QY	235	LTFYLPESPRLVSKGRMAEAKVLOKRGKDVGEISLLEGEVGGDTISIEEYIG	294
Db	178	LTFYLPESPRLVSKGRMAEAKVLOKRGREDYIDEMALLVEGLDIGEKTMEDLLVT	237
QY	235	PATEAADLVYTGDKK-QITLYPBEQCSYIAPSPSGPIMLGSVLSLASHHGMVNQSV	353
Db	238	LEDHGGDITLFTYDDEBQWRLYGTHTNOSYIAPVEEQ--NSLSGRSHRHSGLANQSMI	294
QY	354	LMDDPIVTLGSHVHNNPQAGGSMRSTLPFGSMFVTDQ--HAKNQOMQ---EENLHRD	408
Db	295	LMDDPIVTLGSHLKKKPEAGNTRSGIFPHFGSMFTTADAPGKPAHMKDIESHNKD	354
QY	409	DEEVSADGAGDYED--NLHSLGRQATGAEGKIIVHHGHSALSMRQTLIGEGD	465
Db	355	NDXYATDDGAGDDDBSDNDLRSLMRQTTSMQ-KMIIHPITSGTILSKRRHSITLMQ-GN	412
QY	466	GVSSTDIQGGWOLAMKMSKEGNGRKEGGFKRYVLIHQEGVPCSGRGSIVSLDGGADVFE	525
Db	413	GGSSMGIGGGMHMGHYENDE-----YKRYLKEPDAB-SRGSITISIPGPD--G	460
QY	526	GSEFHYAALVYQSALFSKGLAEPRNSDAAMHSEVAAGGSWKDLPFGVRRALLVGY	585
Db	461	GGSYTHASALVRSVLYGPKS-----YHGSAMVPEKXIASGPMSALLBPGVRRALVGV	515
QY	586	GIQIQOQFPGINGVLYYTPQILEQAGVAVLISFKGSSASASITLISLTMLPCIGFA	645
Db	516	GIQIQOQFPGINGVLYYTPQILERAGVDIILSSLGSSISASBLISGLTTLMLPAIYVA	575
QY	646	MLMLDLSGRRLPLGTIPILIASLVILVAVNSLIDCTLAHALHSTVSVIYFCCPFWGFQ	705
Db	576	MLMLDLSGRRLPLGTIPILIASLVILVAVLISELHIEKYNNAALSTGCVALYFCEFFWGYG	635
QY	706	PIPNILCAEIPTRVRGCLCAICTFTWIGDIIIVYSLPVMALAGLAVGSIVAVCLI	765
Db	636	PIPNILCSEIPTRVAGCLCAICAMFWIGDIIIVYSLPVLSSIGLVGFSIVAAVCVI	695
QY	766	SFVYFLKVPETKMPLEVITEFENAVGAAQAQA	800
Db	696	SWIFVYMKVPETKMPLEVITDPAQGA-QAQA	729

RESULT 7	
ABB92622	
ID ABB92622	standard; Protein, 729 AA.
XX	
AC ABB92622;	
XX	
DT 31-MAY-2002	(first entry)
XX	
DE Herbicidially	active polypeptide SEQ ID NO 1833.
XX	
KW Herbicidal, plant; agriculture; herbicide.	
XX	
OS Arabidopsis thaliana.	
XX	
PN WO200210210-A2.	
XX	
ED 07-FEB-2002.	
XX	
XX 28-AUG-2001;	2001WO-EP09892.
PF 28-AUG-2001;	2001WO-EP09892.
XX	
PR 28-AUG-2001;	2001WO-EP09892.
XX	
PA (FARB) BAYER AG.	

XX	Tietjen K, Weidler M;
PI	WPI; 2002-269010/31.
DR	
XX	
XX	Identifying plant target proteins for herbicidally active compounds,
PT	comparing aligning and comparing nucleic acid or amino acid sequences
PT	from plant with nucleic acid or amino acid sequences from non-plant
PT	organisms -
XX	
PS	Claim 5; SEQ ID NO 1833; 261pp + Sequence Listing; English.
XX	
CC	The invention relates to identifying target proteins
CC	(AB90790-AB94016) for herbicidally active compounds, comprising
CC	aligning and comparing nucleic acid or amino acid sequences from plant
CC	with nucleic acid or amino acid sequences from non-plant organisms using
CC	suitable search parameters, where plant sequences having an E-value
CC	greater by a factor of 3 than the E-value of most similar non-plant
CC	sequences are selected. The polypeptides or nucleic acids encoding them
CC	are useful for identifying modulators. The identified modulators are
CC	useful as herbicides.
SQ	Sequence 729 AA;
Query Match	55.1%; Score 2245; DB 23; Length 729;
Best Local Similarity	60.4%; Pred.No.9,5e-208;
Matches	457; Conservative 110; Mismatches 144; Indels 46; Gaps 14
OY	MSGAVLVIVASIGMLTLOGMDNATTAANVLTYKKEFOLNEPTEGLIVSMISLGATVYT 115
Db	1 MRSVTVLVAALAAIGMLOGMDNATTAGAVIYIKKEFHLEKEPKITELIVAMSLIGATLT 60
OY	116 TFGSGPLSDSIGRRPMILISSIIYFPFSGIMTMSPNVVYLARFYDGFGIGLATVLPVLY 175
Db	61 TFGSGVSPKVGRRSMTLLSSVLYFPSSIVMFMSPNVVYLLFRLLIDGFGIGLATVLPVIY 120
OY	176 ISETAPSRIRGLINLTLPQFSGGGMFLSYCWMPFGMSLSSPPWRIMLVATPLSLEFFGL 235
Db	121 ISETAPSRIRGLINLTLPQFCGGGMFLSYCLVFNGMSLDESPSWRLMLGVLSIPSIAPVYL 180
OY	236 TIIFYPEGPRLMWISKGRAAKVYLQKRGKDVSGETSLILEGLEVGADTSIEEYTIOP 295
Db	181 AAFPIPSRPRLMWISKGRNDKARQVYLQRKRGDVSGELALLVEGLGVGDTSTEEVYIER 240
OY	296 ATGAADDLVTDGDKEOITLYGPEGQSQSIARPSKCPIMLGSVLSSARRHGSVMNQSVPLM 355
Db	241 DNEEHGEGNELPRKQDIKLPGEDGQSNMAKVPKGO---SIALASRQGSMUPRGCSIM 296
OY	356 DPITYTLPESSVENNP--OAGSMRETLPPNFGSMFSVTDQHAKNDQMDENLHRDBEYA 413
Db	297 DPLVTYLFPSIHENLPSENMMNASRSEMLEPPNMGISLIGMGR--QESQWDE---RNND-- 349
OY	414 SDAGAGVDIEDNLSPILSRQATGAAGXOIIVNHGHRSALSMSRQTL-LDGEGDGVGSTDI 472
Db	350 ---SSDDENINSIFLSPOTT--EPDD--YHQRTVGTMRHQSSIFPMANVGETATATSI 401
OY	473 GGGNQLAMKWEKEEENGK--EGGFKRYVLHO-----GYVGSRRGSIVSL-PCGGDV 523
Db	402 GGGWGLAMKYNKYGADGKRVNGGLGRMYIHETANNNTNNIPFSRGSLSPHPREGDH 461
OY	524 FEGSEFEVDAALVQSOLFSGLAEPNSDAAMPHPSEVAAKGRWKDLPEFGVRRALIV 583
Db	462 DVNVGVYDAALVQSASMPGCKGETAML-----PKEV-KDGPGRBELKEGVVKALMV 514
OY	584 GVGIOTILOOFAGINCVALYTPQIIIEQAQAVAVILSKFGISSASAILISLTLLTMPCIG 643
Db	515 GVGIOIILQDFAGINCVMYTPQIIIEBTGVSLLTNMGISAESASILISALTLLMPCIC- 573
OY	644 FAMLLMDISGRFLLGTPIILIASVILVSNLIDLGLAALASTVSVIYVCFCEFWG 703
Db	574 -----LVSKRSLSMLSTIPIILISLVTLVIGLSVNLGGSINALISTASVYYLSCFWG 626
OY	704 FGPIENLICAEIPTRVNGCIAICFTFWIGDIIIVTYSPLYMLNAIGLAGVSIYAVVC 763

```

Db      627  FGAIFNIIICSELPFISVAGLCITTCALTFWICDIIIVTYTLPLPMASIGIAGVFGLIAYVC 686
Oy      764  LISFVFELKVPEIKMPLEVIITEFPFVAVGAKQAANA 800
Db      687  AVAMVFYLIKVPETKGMPELVISEFFSVGAKQADAA 723

RESULT 8
AAU97205
ID      AAU97205 standard; Protein; 486 AA.
AC      AAU97205;
XX      27-AUG-2002 (first entry)
DT      Portion of a soybean sugar transport protein encoded by ssl.pk0022.fl.
DE      Soybean; Arabidopsis thaliana-like sugar transport protein;
KW      carbohydrate transport; grain filling; annual field crop;
XX      plant.
XX      Glycine max.
OS      Glycine max.
PN      US6383776-B1.
XX      07-MAY-2002.
XX      14-APR-1999; 99US-0291922.
XX      24-APR-1998; 98US-083044P.
PR      (DUPLO ) DU PONT DE NEMOURS & CO E I.
XX      Allen SM, Hiltz WD, Kinney AJ, Tingey SV;
XX      MPI: 2002-451386/48.
XX      N-PSDB; ARK51966.
DR      New nucleic acid encoding plant sugar-transport proteins, useful for
PT      preparing transgenic plants with altered carbohydrate distribution -
PT      Claim 1; Fig 1; 54pp; English.
XX      The present invention relates to the isolation of plant polynucleotide
CC      sequences encoding an Arabidopsis thaliana-like sugar transport
CC      protein or Beta vulgaris-like sugar transport protein. The
CC      polynucleotide sequences are useful for altering the level of sugar
CC      transport proteins in plants, i.e. for control of carbohydrate transport
CC      and distribution in plant cells, e.g. during grain filling of annual
CC      field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
CC      carbohydrate flows and sugar transport. The polynucleotide
CC      sequences can also be used to isolate cDNA sequences and genes that
CC      encode homologues of the new proteins. The present sequence
CC      represents a portion of a soybean Arabidopsis thaliana-like sugar
CC      transport protein.
XX      Sequence 486 AA;
SQ
Query Match 41.0%; Score 1670.5; DB 23; Length 486;
Best Local Similarity 68.1%; Pred. No. 2.2e-152;
Matches 333; Conservative 57; Mismatches 84; Indels 15; Gaps 8
Oy      308  DKEQITLLGPERGGSWIARPSKGPITMLGSVLSIARHSGMVNQSVPLMDPIVTLFGSVHE 367
Db      5 EKDQIKLKGPEGGGQSWARPAVAGPNSVG---LVRKKSMMNPS-SYVDPIVTLFGSVHE 59
Oy      368  NMPQAGGSMRSTLFPNFGSMESVYTDQAKNEQWDEENLRHDEERYASDQAGDYEDNTHS 427
Db      60 KLPEFG---STLFPFGSMFVSFGNQNPNEDWDEESLAREGDDYVSD--AGDSDDNTHS 113
Oy      428  PLISQATGASGEKDIIVHGHRGSAISMRQTLT--GEGDGVNSTIIGGQWLAKWKSKE 486

```

Db	111	PLLSRQTTSLD-KDIPPHAHSLA-SMGQSLIHENSEPFGSGIGGQMLANKMSERE	171
Qy	487	GENGRKEGGEKRVYIHQEGVPCSGRRGSIVSLPGGDVEEGSEFYHAALVQSALFSGKL	546
Db	172	GPQKKEGGEKRYIHLQDGGSGRRGSVSLP-GGDLPTDSBVAQAAALVQSPALYNEDL	230
Qy	547	AERPMEDAMVHSEVFAAKGSMWCOLPEFGVRALLVGVGQIIQQPFGINGVYIYPIQI	606
Db	231	MRQRPVGPMAIHSEETIAKPSMSDLPEFGVGHALIVGVGMQIIQQPFGINGVYIYPIQI	290
Qy	607	LEQAGVAVILSKKGLSSASALISLTLTLMPLCTIGFAMLLMPLDGGRRFLLGTIPILI	666
Db	291	LEQAGVGYLLSLGLGSTSSFLISAVTTLMLPCTIATAMLMPLDSGRRTLLSTIYVLI	350
Qy	667	ASVLIVLNSLIDLTGLTALLSTVSVIYVPCCFWGGFPIPNILCAEIPETRVRGLCIA	726
Db	351	AALLIVLGLSLVDLGTANASISTISVYVFCFFWGGFPIPNILCAEIPETRVRGLCIA	410
Qy	727	ICAFPTWIGDIIVTYSLLPVMNLALIGLAVFSTIYAVVCLISVFPFLKVPETKGMPLFVIT	786
Db	411	ICALTITWIDIIIVTYTTLPMNLNSVGLAGVFGIYAVVCFIAMVFPFLKVPETKGMPLFVIT	470
Qy	787	EPFVAGAKQ 795	
Db	471	EPFVAGAKQ 479	
RESULT 9			
ABU08330			
ID	ABU08330	standard; Protein; 486 AA.	
AC	ABU08330;		
XX			
DT	29-MAY-2003	(first entry)	
XX			
DE	Soybean sugar transport protein #2.		
XX			
KW	Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;		
KW	plant sugar transport protein; carbohydrate transport; soybean;		
KW	carbohydrate distribution; plant.		
XX			
OS	Glycine max.		
XX			
PN	US2002178468-A1.		
PD			
XX	28-NOV-2002.		
XX			
PE	17-JAN-2002; 2002US-0051902.		
XX			
PR	24-APR-1998; 98US-083044P.		
PR	14-APR-1999; 99US-0291922.		
XX			
PA	(ALTE/) ALLEN S M.		
PA	(HITZ/) HITZ W D.		
PA	(KINN/) KINNEY A J.		
PA	(TING/) TINGEY S V.		
XX			
PI	Allen SM, Hitz WD, Kinney AJ, Tingey SV;		
XX			
DR	WPI: 2003-340957/32.		
DR	N-PSDB; ABA93202.		
XX			
PT	Novel plant sugar transport proteins and nucleic acid encoding the		
PT	protein useful for producing transgenic plants having altered levels of		
PT	sugar transport protein _		
XX			
PS	Claim 5; Fig 1; 56pp; English.		
XX			
CC	The present invention relates to the isolation of Arabidopsis		
CC	thaliana-like or Beta vulgaris-like sugar transport proteins, and		
CC	the polynucleotide sequences encoding them. The plant sugar transport		
CC	proteins of the invention have been isolated from corn, rice, soybean,		
CC	and wheat. The polypeptides of the invention may be used for altering		

CC the level of expression of a sugar transport protein in a host cell,
 CC by transforming a host cell with a chimeric construct encoding all,
 CC or a portion of the sugar transport protein, in sense or antisense
 CC orientation. Particularly, the polypeptides may provide a means to
 CC control carbohydrate transport and distribution in plants.
 CC ABU08326-ABU08333 represent Arabidopsis thaliana-like sugar
 CC transport proteins.

XX Sequence 486 AA;

Query Match 41.0%; Score 1670.5; DB 24; Length 486;
 Best Local Similarity 68.1%; Pred. No. 2.2e-152;
 Matches 333; Conservative 57; Mismatches 84; Indels 15; Gaps 8;

QY 308 DKEDTLVGPESGSMWAPSPKPIMLGSLVSLASHRSMNQSPLMDPIVTLFGSVHE 367
 DB 5 EKDQIKLGPESGSMWAPSPKPIMLGSLVSLASHRSMNQSPLMDPIVTLFGSVHE 59
 QY 368 NMPQAGSMRSTLFPNFGSMFSTVDQAKNEQWDEENLRDDEYASDAGGDEYEDNLHS 427
 DB 60 KLPEFG---STLFPFGSMFSGVGNQPRNEDMDESLAREDDVYSD--AGDSNDNQ 113
 QY 428 PLLRQATGAEGKDIHHGHRGSLSMRQTL--GEGDGVSTDIIGGQWOLAMKSEKE 486
 DB 114 PLISRQTTSLD-KDIPRAHSLA-SMRQSLHNSGEPGSGTIGGQWOLAMKSEKE 171
 QY 487 GENGKREGFPRVYHOGVSGSRGSIYSLPGGDVSESGEPVHAALVQSALFSKGL 546
 DB 172 GPDEKREGFPRVYHOGVSGSRGSIYSLPGGDVSESGEPVHAALVQSALFSKGL 230
 QY 547 AEPMSDAMVHPSEVAKGRMKDLFEPGVRAALLVGVGIQIQQFAGINGVLYTPOI 606
 DB 231 MRQSPVGNAMHPSSTIAKGRMSDLFEPGVRAALLVGVGIQIQQFAGINGVLYTPOI 290
 QY 607 LEQAGVAIVISKFGLSASASILLSTLTLLMPCIGFAMLLMDISGRFLLGTIPILI 666
 DB 291 LEQAGVGYLLSLGISTSSSFLISAVTLLMLPCIAIAMRLMDISGRFLLGTIPILI 350
 QY 667 ASLIVIVSNLIDLTALHALLSTVSVYVFCFPMGSGPIPNLICATIPPRVGLCIA 726
 DB 351 AALLIVGLSVLDLSTANASISTISVYVFCFPMGSGPIPNLICATIPPRVGLCIA 410
 QY 727 ICAFFFWIGDIIVTVSLPVMNLAIAGVFSIYAVCLISFVFLAKVETKGMPLLEYIT 786
 DB 411 ICALTFWICDIIVTVSLPVMNLAVGLAGVFIYAVVCIYAVFVLKVPETKGMPLLEYIT 470
 QY 787 EFPVAGAKQ 795
 DB 471 EFPVAGAKQ 479

RESULT 10

AAU97207
 ID AAU97207 standard; Protein; 345 AA.

AC AAU97207;

DT 27-AUG-2002 (first entry)

DE Portion of a wheat sugar transport protein encoded by a contig.

XX Wheat; Arabidopsis thaliana-like sugar transport protein;
 KW carbohydrate transport; grain filling; annual field crop;
 KM plant.

OS Triticum aestivum.

PN US6383776-B1;

XX 07-MAY-2002.

PF 14-APR-1999; 99US-0291922.

XX

PR 24-APR-1998; 98US-083044P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Allen SM, Hiltz WD, Kinney AJ, Tingey SV;

XX WPI; 2002-451386/48.

DR N-PSDB; ABK51968.

PT New nucleic acid encoding plant sugar-transport proteins, useful for
 XX preparing transgenic plants with altered carbohydrate distribution

XX Claim 1; Fig 1; 54pp; English.

CC The present invention relates to the isolation of plant polynucleotide
 CC sequences encoding an Arabidopsis thaliana-like sugar transport
 CC protein or Beta vulgaris-like sugar transport protein. The
 CC polynucleotide sequences are useful for altering the level of sugar
 CC transport proteins in plants, i.e. for control of carbohydrate transport
 CC and distribution in plant cells, e.g. during grain filling of annual
 CC field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
 CC carbohydrate flows and sugar transport. The polynucleotide
 CC sequences can also be used to isolate cDNA sequences and genes that
 CC encode homologues of the new proteins. The present sequence
 CC represents a portion of a wheat Arabidopsis thaliana-like sugar
 CC transport protein.

XX Sequence 345 AA;

Query Match 34.9%; Score 1424; DB 23; Length 345;
 Best Local Similarity 82.9%; Pred. No. 9.2e-129;
 Matches 281; Conservative 24; Mismatches 32; Indels 2; Gaps 2;

QY 462 EGDGVSSTDIIGGQWOLAMKSEKEGKRGKRVYHOGVPSRGSIVSLPGCG 521
 DB 4 EGGEAVSSTGIGGQWOLAMKSEKEDGKKEGFKRIVHOGVADSRGSIVSLPGCG 63
 QY 522 DVPE-GSEPFHAALVQSALFSKGLAEPH--SDAAVHPSEVAKGRMKDLFEPGVR 579
 DB 64 DATGSGGFTHAALVSHSALYSKDLMEERAAQPAWHPUEAPKGSIMKDLFEPGVR 123
 QY 580 ALTVGVGIQIQQFAGINGVLYTPOILEQAGVAIVLSKFGLSASASILLSTLTLLML 639
 DB 124 ALFVGVGIQIQQFAGINGVLYTPOILEQAGVAIVLSNGLSSASASILLSTLTLLML 183
 QY 640 PCIGFAMLLMDISGRFLLGTIPILIASVIVLSNLDLGTALHALLSTVSVYVFC 699
 DB 184 PSIGVAMRLMDISGRFLLGTIPILIASVIVLGVVNVINLSTVPHAVLSTVSVYVFC 243
 QY 700 FVWGFGPIPNLICATIPPRVGLCIAFTFWIGDIIVTVSLPVMNLAIAGVFSIY 759
 DB 244 FVWGFGPIPNLICATIPPRVGLCIAFTFWIGDIIVTVSLPVMNLAIAGVFSIY 303
 QY 760 AVVCLISFVFLKVPETKGMPLVITEFPAVGAQQA 798
 DB 304 AVVCLISFVFLKVPETKGMPLVITEFPAVGAQQA 342

RESULT 11

ABU08332
 ID ABU08332 standard; Protein; 345 AA.

AC ABU08332;

DT 29-MAY-2003 (first entry)

DE Wheat sugar transport protein #2.

XX Arabidopsis thaliana-like sugar transport protein; corn, rice, wheat;
 KW plant sugar transport protein; carbohydrate transport; soybean;
 KM carbohydrate distribution; plant.

OS Triticum aestivum.

XX US2002178468-A1.
 PN
 XX 28-NOV-2002.
 PD
 XX 17-JAN-2002; 2002US-0051902.
 PF
 XX 24-APR-1998; 98US-083044P.
 PR
 XX 14-APR-1999; 99US-0291922.
 XX
 PA (ALLE/) ALLEN S M.
 PA (HITZ/) HITZ W D.
 PA (KINNEY/) KINNEY A J.
 PA (TINGEY/) TINGEY S V.
 XX
 PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
 DR WPI; 2003-340957/32.
 DR N-PSDB; ABX93204.
 XX
 PT Novel plant sugar transport proteins and nucleic acid encoding the
 PT protein useful for producing transgenic plants having altered levels of
 PT sugar transport protein
 PS
 XX Claim 5; Fig 1; 56pp; English.
 XX
 CC The present invention relates to the isolation of Arabidopsis
 CC thaliana-like or Beta vulgaris-like sugar transport proteins, and
 CC the polynucleotide sequences encoding them. The plant sugar transport
 CC proteins of the invention have been isolated from corn, rice, soybean,
 CC and wheat. The polypeptides of the invention may be used for altering
 CC the level of expression of a sugar transport protein in a host cell,
 CC by transforming a host cell with a chimeric construct encoding all,
 CC or a portion of the sugar transport protein, in sense or antisense
 CC orientation. Particularly, the polypeptides may provide a means to
 CC control carbohydrate transport and distribution in plants.
 CC ABU08336-ABU08333 represent Arabidopsis thaliana-like sugar
 CC transport proteins.
 CC
 SQ Sequence 345 AA;
 Query Match 34.9%; Score 1424; DB 24; Length 345;
 Best Local Similarity 82.9%; Pred. No. 9.2e-129;
 Matches 281; Conservative 24; Mismatches 32; Indels 2; Gaps 2;
 QY 462 EGGDGVSSSTIDIGGQQLAMKSEKGGKRGKRRVYLHQEGVPGSRGIVSLPGGG 521
 DB 4 EGGGAVSSSTIGIGGQQLAMKSEKGGKRGKRRVYLHQEGVADSRGIVSLPGGG 63
 QY 522 DVFE-GSEFVHAALVSSALFSGKLAERPM-SDAMVPSVAKKGSRWKLPFRGVVR 579
 DB 64 DVTQGGSGGTHAALVSHALYSKDLMEERMAAGPMHPLAAPRGSTWKLPFRGVVR 123
 QY 580 ALLVGVGIQILQOPAGINGVLYTTPQILEQAGVAVILSKFGLSSASASLISLTLLML 639
 DB 124 ALFVGVGIQILQOPAGINGVLYTTPQILEQAGVAVILSKFGLSSASASLISLTLLML 183
 QY 640 PCIGPAMLMDISGRFLLGTPTLILASLVILVNSLIDLTALHALLISTVSVIYFCC 699
 DB 184 PSIGVAMRLMDISGRFLLGTPTLILASLVILVNSLIDLTALHALLISTVSVIYFCC 243
 QY 700 FPMGSGPIPNILCAEIPFRVRGGLCAICAFPMIGDIIIVTSLPMLNAGLAVFSY 759
 DB 244 FPMGSGPIPNILCAEIPFRVRGGLCAICAFPMIGDIIIVTSLPMLNAGLAVFSY 303
 QY 760 AVVCLISFVFLKVPETKGMPLVETFEFAVGAQAAA 798
 DB 304 AVVCCIAFVFLKVPETKGMPLVETFEFAVGAQAAA 342

RESULT 12
 AAU97208
 ID AAU97208 standard; Protein; 228 AA.

XX AAU97208;
 AC
 XX 27-AUG-2002 (first entry)
 DT
 XX Portion of a wheat sugar transport protein encoded by wrein.pk0006.b4.
 DE
 XX Wheat; Arabidopsis thaliana-like sugar transport protein;
 KW carbohydrate transport; grain filling; annual field crop;
 KW plant.
 OS
 XX Triticum aestivum.
 PN
 XX US6383776-B1.
 PD
 XX 07-MAY-2002.
 XX
 XX 14-APR-1999; 99US-0291922.
 XX
 XX 24-APR-1998; 98US-083044P.
 PR
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 PA
 XX Allen SM, Hitz WD, Kinney AJ, Tingey SV;
 PI
 XX WPI; 2002-451386/48.
 DR
 XX N-PSDB; ABK51969.
 XX
 PT New nucleic acid encoding plant sugar-transport proteins, useful for
 PT preparing transgenic plants with altered carbohydrate distribution
 PS
 XX Claim 1; Fig 1; 54pp; English.
 XX
 CC The present invention relates to the isolation of plant polynucleotide
 CC sequences encoding an Arabidopsis thaliana-like sugar transport
 CC protein or Beta vulgaris-like sugar transport protein. The
 CC polynucleotide sequences are useful for altering the level of sugar
 CC transport proteins in plants, i.e. for control of carbohydrate transport
 CC and distribution in plant cells, e.g. during grain filling of annual
 CC field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
 CC carbohydrate flows and sugar transport. The polynucleotide
 CC sequences can also be used to isolate cDNA sequences and genes that
 CC encode homologues of the new proteins. The present sequence
 CC represents a portion of a wheat Arabidopsis thaliana-like sugar
 CC transport protein.
 CC
 SQ Sequence 228 AA;
 Query Match 23.1%; Score 943; DB 23; Length 228;
 Best Local Similarity 80.6%; Pred. No. 1.8e-82;
 Matches 179; Conservative 25; Mismatches 18; Indels 0; Gaps 0;
 QY 574 EPGVRRALLVGVGIQILQOPAGINGVLYTTPQILEQAGVAVILSKFGLSSASASLISLT 633
 DB 1 EPGVRRALLVGVGIQILQOPAGINGVLYTTPQILEQAGVAVILSKFGLSSASASLISLT 60
 QY 634 TTLMLPCIGPAMLMDISGRFLLGTPTLILASLVILVNSLIDLTALHALLISTVSV 693
 DB 61 TTLMLPCIGPAMLMDISGRFLLGTPTLILASLVILVNSLIDLTALHALLISTVSV 120
 QY 694 IYVPCFPMGSGPIPNILCAEIPFRVRGGLCAICAFPMIGDIIIVTSLPMLNAGL 753
 DB 121 IYVPCFPMGSGPIPNILCAEIPFRVRGGLCAICAFPMIGDIIIVTSLPMLNAGL 180
 QY 754 GVFSIYAVVCLISFVFLKVPETKGMPLVETFEFAVGAQO 795
 DB 181 GVFSIYAVVCLISFVFLKVPETKGMPLVETFEFAVGAQO 222

RESULT 13
 ABU08333
 ID ABU08333 standard; Protein; 228 AA.

AC ABU08333;
 XX 29-MAY-2003 (first entry)
 DT Wheat sugar transport protein #3.
 XX Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
 KM plant sugar transport protein; carbohydrate transport; soybean;
 KM carbohydrate distribution; plant.
 XX Trifolium aestivum.
 OS
 XX US2002178468-A1.
 PN
 XX 28-NOV-2002
 PD
 XX 17-JAN-2002; 2002US-0051902.
 PF
 XX 24-APR-1998; 98US-083044P.
 PR 14-APR-1999; 99US-0291922.
 XX
 XX (ALDE/) ALLEN S M.
 PA (HITZ/) HITZ W D.
 PA (KINNEY/) KINNEY A J.
 PA (TING/) TINGEY S V.
 XX
 PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
 XX
 DR WPI; 2003-340957/32.
 XX
 DR N-PSDB; ABX93205.
 XX
 PT Novel plant sugar transport proteins and nucleic acid encoding the
 PT protein useful for producing transgenic plants having altered levels of
 PT sugar transport protein -
 XX
 PS Claim 5; Fig 1; 56pp; English.
 XX
 CC The present invention relates to the isolation of Arabidopsis
 CC thaliana-like or Beta vulgaris-like sugar transport proteins, and
 CC the polynucleotide sequences encoding them. The plant sugar transport
 CC proteins of the invention have been isolated from corn, rice, soybean,
 CC and wheat. The polypeptides of the invention may be used for altering
 CC the level of expression of a sugar transport protein in a host cell,
 CC by transforming a host cell with a chimeric construct encoding all,
 CC or a portion of the sugar transport protein, in sense or antisense
 CC orientation. Particularly, the polypeptides may provide a means to
 CC control carbohydrate transport and distribution in plants.
 CC ABU08326-ABU08333 represent Arabidopsis thaliana-like sugar
 CC transport proteins.
 CC
 XX
 SQ Sequence 228 AA;
 Query Match 23.1%; Score 943; DB 24; Length 228;
 Best Local Similarity 80.6%; Pred. No. 1.8e-82;
 Matches 179; Conservative 25; Mismatches 18; Indels 0; Gaps 0;

QY 574 EGVRRALLVGVGIQLQFAGINGVLYTTPQILEQAGVAVILSKRGSSASASILISGL 633
 DB 1 EGVGHALFVIGLQILOFAGINGVLYTTPQILEQAGVAVILSNIGSSSSASILISGL 60
 QY 634 TTLMLPCTGFMALMDLSGRFLLGTPILIASIVIVVNSLIDGLTALALSTSYV 693
 DB 61 TTLMLPSTIGIMRMDMSGRFLLSTIPVILVALAVLAVLDVGMVAALSTISV 120
 QY 694 IYVPCFVWGFPIPIILCAEIFFPVRGLCIAICAFPMIGDIITYTSLPMLNAGLA 753
 DB 121 IYVPCFVWGFPIPIILCAEIFFPVSVRGICIAICLTFWIGDIITYTSLPMLNAGLA 180
 QY 754 GVFSIYAVVGLISFVFLKVPETKMPLEVTTEFFAVGAKQ 795
 DB 181 GVFGIYAVCVLAIFVVMKVPETKMPLEVTTEFFSVGAKQ 222

RESULT 14
 AAU97203
 ID AAU97203 standard; Protein; 131 AA.
 XX
 AC AAU97203;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE Portion of a rice sugar transport protein encoded by rl66.pk0003.d5.
 XX
 KM Rice; Arabidopsis thaliana-like sugar transport protein;
 KM carbohydrate transport; grain filling; annual field crop;
 KM plant.
 XX
 XX Oryza sativa.
 OS
 XX US6383776-B1.
 PN
 XX 07-MAY-2002.
 PD
 XX 14-APR-1999; 99US-0291922.
 PF
 XX 24-APR-1998; 98US-083044P.
 PR
 XX (DUBO) DU PONT DE NEMOURS & CO E I.
 PA
 XX
 PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
 XX
 DR WPI; 2002-451386/48.
 XX
 DR N-PSDB; ABK51964.
 XX
 PT New nucleic acid encoding plant sugar-transport proteins, useful for
 PT preparing transgenic plants with altered carbohydrate distribution -
 XX
 PS Claim 1; Fig 1; 54pp; English.
 XX
 CC The present invention relates to the isolation of plant polynucleotide
 CC sequences encoding an Arabidopsis thaliana-like sugar transport
 CC protein or Beta vulgaris-like sugar transport protein. The
 CC polynucleotide sequences are useful for altering the level of sugar
 CC transport proteins in plants, i.e. for control of carbohydrate transport
 CC and distribution in plant cells, e.g. during grain filling of annual
 CC field crops (e.g. corn, rice, soybeans, and wheat), and for studying
 CC carbohydrate flows and sugar transport. The polynucleotide
 CC sequences can also be used to isolate cDNA sequences and genes that
 CC encode homologues of the new proteins. The present sequence
 CC represents a portion of a rice Arabidopsis thaliana-like sugar
 CC transport protein.
 CC
 XX
 SQ Sequence 131 AA;
 Query Match 13.8%; Score 562; DB 23; Length 131;
 Best Local Similarity 80.5%; Pred. No. 5.8e-46;
 Matches 103; Conservative 17; Mismatches 8; Indels 0; Gaps 0;

QY 669 LVILVNSLIDGLTALALSTVSYIYFCCFWGFGPIPIILCAEIFFPVRGLCIAIC 728
 DB 1 VLTILVNLIDVGMVHASISTSVIYFCCFWGFGPIPIILCAEIFFPVRGLCIAIC 60
 QY 729 AFTFMIGDIITYTSLPMLNAGLAGVSIYAVVCLISFVFLKVPETKMPLEVTTEF 788
 DB 61 AFTFMIGDIITYTSLPMLNAGLAGVSIYAVVCLISFVFLKVPETKMPLEVTTEF 120
 QY 789 FAVGAKQA 796
 DB 121 FSVGAKQA 128
 RESULT 15
 ABU08328
 ID ABU08328 standard; Protein; 131 AA.
 XX
 AC ABU08328;

```

XX 29-MAY-2003 (first entry)
DT Rice sugar transport protein #2.
XX
DE Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
XX plant sugar transport protein; carbohydrate transport; soybean;
XX carbohydrate distribution; plant.
OS Oryza sativa.
XX
XX US2002178468-A1.
XX
XX 28-NOV-2002.
XX
XX 17-JAN-2002; 2002US-0051902.
XX
XX 24-APR-1998; 98US-083044P.
XX
XX 14-APR-1999; 99US-0291922.
XX
XX (ALLEN/) ALLEN S M.
XX PA (HITZ/) HITZ W D.
XX PA (KINNEY/) KINNEY A J.
XX PA (TINGE/) TINGEY S V.
XX
XX Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
XX WPI: 2003-340957/32.
XX
XX N-PSDB; ABX93200.
XX
XX Novel plant sugar transport proteins and nucleic acid encoding the
XX protein useful for producing transgenic plants having altered levels of
XX sugar transport protein
XX
XX Claim 5; Fig 1; 56pp; English.
XX
XX The present invention relates to the isolation of Arabidopsis
XX thaliana-like or Beta vulgaris-like sugar transport proteins, and
XX the polynucleotide sequences encoding them. The plant sugar transport
XX proteins of the invention have been isolated from corn, rice, soybean,
XX and wheat. The polypeptides of the invention may be used for altering
XX the level of expression of a sugar transport protein in a host cell,
XX by transforming a host cell with a chimeric construct encoding all,
XX or a portion of the sugar transport protein, in sense or antisense
XX orientation. Particularly, the polypeptides may provide a means to
XX control carbohydrate transport and distribution in plants.
XX
XX ABU08326-ABU08333 represent Arabidopsis thaliana-like sugar
XX transport proteins.
XX
XX
SQ Sequence 131 AA;

```

Query Match 13.8%; Score 562; DB 24; Length 131;
Best local Similarity 80.5%; Pred. No. 5.8e-46;
Matches 103; Conservative 17; Mismatches 8; Indels 0; Gaps 0;

```

QY 669 LVTLVSNLIDIGTTLAHLISVSVIVYFCFVWGFGPIPNILCAEIPTRVGLCIATC 728
DB 1 VLTLLIVNIDVGTWTHASISTSVILYRCFFVWGFGPIPNILCAEIPTRVGLCIATC 60
QY 729 APTFWIGDIIVTVYSLPVMNNAIGLGVFSIYAVVCLISFVFLKVPETKGMPLVEVITEF 788
DB 61 ALTFWIGDIIVTVYTLVPMNNAIGLGVFSIYAVVCLISFVFLKVPETKGMPLVEVITEF 120
QY 789 FAVGAKQA 796
DB 121 FSVGAKQA 128

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 18:54:10 ; Search time 30 Seconds

(without alignments)
1128.290 Million cell updates/sec

Title: US-10-051-909-32

Perfect score: 4075
Sequence: 1 IRSSGWLAVQPFPTDLDR.....PLEVITPEFVAGKQAANKA 800

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep: *
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep: *
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep: *
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep: *
5: /cgn2_6/ptodata/2/1aa/PTCUS_COMB.pep: *
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3517	86.3	747	4	US-09-291-922-2 Sequence 2, Appl1
2	2674	65.6	737	4	US-09-291-922-8 Sequence 8, Appl1
3	2584	63.4	729	4	US-09-291-922-29 Sequence 29, Appl1
4	1670.5	41.0	486	4	US-09-291-922-10 Sequence 10, Appl1
5	1424	34.9	345	4	US-09-291-922-14 Sequence 14, Appl1
6	943	23.1	228	4	US-09-291-922-16 Sequence 16, Appl1
7	562	13.8	131	4	US-09-291-922-6 Sequence 6, Appl1
8	496	12.2	523	4	US-09-291-922-24 Sequence 24, Appl1
9	489.5	12.0	513	4	US-09-291-922-20 Sequence 20, Appl1
10	489	12.0	529	4	US-09-291-922-28 Sequence 28, Appl1
11	488	12.0	549	4	US-09-291-922-30 Sequence 30, Appl1
12	485	11.9	510	4	US-09-291-922-22 Sequence 22, Appl1
13	463.5	11.4	539	4	US-09-291-922-26 Sequence 26, Appl1
14	421.5	10.3	584	2	US-08-928-692-13 Sequence 13, Appl1
15	421.5	10.3	584	4	US-09-339-972-13 Sequence 13, Appl1
16	384.5	9.4	534	2	US-09-031-392-4 Sequence 4, Appl1
17	384.5	9.4	534	3	US-09-299-549-4 Sequence 4, Appl1
18	384.5	9.4	534	4	US-09-610-417-4 Sequence 4, Appl1
19	365	9.0	493	2	US-09-031-392-10 Sequence 10, Appl1
20	365	9.0	493	3	US-09-299-549-10 Sequence 10, Appl1
21	365	9.0	493	4	US-09-610-417-10 Sequence 10, Appl1
22	358	8.8	488	2	US-08-928-692-10 Sequence 10, Appl1
23	358	8.8	488	4	US-09-339-972-10 Sequence 10, Appl1
24	346	8.5	524	2	US-08-928-692-12 Sequence 12, Appl1
25	346	8.5	524	4	US-09-339-972-12 Sequence 12, Appl1
26	344	8.4	509	2	US-09-031-392-6 Sequence 6, Appl1
27	344	8.4	509	3	US-09-299-549-6 Sequence 6, Appl1

28	344	8.4	509	4	US-09-610-417-6 Sequence 6, Appl1
29	331	8.1	492	2	US-08-355-844-3 Sequence 3, Appl1
30	331	8.1	492	5	PCT-US95-16126-3 Sequence 3, Appl1
31	330	8.1	488	2	US-08-928-692-11 Sequence 11, Appl1
32	330	8.1	488	4	US-09-339-972-11 Sequence 11, Appl1
33	310	7.6	494	2	US-09-031-392-5 Sequence 5, Appl1
34	310	7.6	494	3	US-09-299-549-5 Sequence 5, Appl1
35	310	7.6	494	4	US-09-610-417-5 Sequence 5, Appl1
36	306.5	7.5	500	2	US-09-031-392-7 Sequence 7, Appl1
37	306.5	7.5	500	3	US-09-299-549-7 Sequence 7, Appl1
38	306.5	7.5	500	4	US-09-610-417-7 Sequence 7, Appl1
39	293	7.2	383	2	US-09-031-392-3 Sequence 3, Appl1
40	293	7.2	383	3	US-09-299-549-3 Sequence 3, Appl1
41	293	7.2	383	4	US-09-610-417-3 Sequence 3, Appl1
42	275.5	6.8	323	4	US-09-134-001C-4635 Sequence 4635, Ap
43	265.5	6.5	475	4	US-09-328-352-5706 Sequence 5706, Ap
44	256.5	6.3	131	4	US-09-291-922-4 Sequence 4, Appl1
45	256	6.3	305	4	US-09-107-532A-4215 Sequence 4215, Ap

ALIGNMENTS

RESULT 1
US-09-291-922-2
Sequence 2, Application US/09291922
Patent No. 6,382,226
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Hitz, Bill
APPLICANT: Kinney, Tony
APPLICANT: Tinney, Scott
TITLE OF INVENTION: Plant Sugar Transport Proteins
FILE REFERENCE: BB-1153
CURRENT APPLICATION NUMBER: US/09/291,922
CURRENT FILING DATE: 1999-04-14
EARLIER APPLICATION NUMBER: 60/083,044
EARLIER FILING DATE: April 24, 1998
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Microsoft Office 97
SEQ ID NO 2
LENGTH: 747
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
NAME/KEY: UNSURE
LOCATION: (129)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (133)..(134)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (144)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (178)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (207)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (218)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (220)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (236)
US-09-291-922-2
Query Match 86.3%; Score 3517; DB 4; Length 747;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 684; Conservative 27; Mismatches 34; Indels 2; Gaps 2;

*Applicant
Priority document*

```

QY 56 MSGAVLVAIVASIGNLQGMNATTAANLYIKKEFOLONEPTVEGLIVSMGLIGATVIT 115
DB 1 MCGAVVVAIAASIGNLQGMNATTAAGAVLYIKKEFNLOSEPLIRGLIAAMFLIGATVIT 60
QY 116 TFSGLSDSIGRRPMLIISIIYFSGILMWSPNVYVLLARFVDPGFGIGLAATLVPLY 175
DB 61 TSPGRADCVGRPMVAVSAVLYFVSGVLMVAPVYIILLAKLIDGFGIGLAATLVPLY 120
QY 176 ISEIAPSEIR-GILNLTLPQFSG-SGGMFLSYCMVFGMSLSPSPDMRIMLGVALIPSLFFPGL 233
DB 121 ISEIAPSEIRSGINLTLPQFSGSGGMFLSYCMVFGMSLSPSPDMRIMLGVALIPSLFFPGL 180
QY 234 GILTYLPSPSPRWLVSKGRMAEAKKYLQRLRGKDVSGELSLLEGLVEGDTISIEBYI 293
DB 181 GILTYLPSPSPRWLVSKGRMAEAKKYLQRLRGKDVSGELSLLEGLVEGDTISIEBYI 240
QY 294 GPATEAADLVTDGDEKQITLYGPBEGSGMIARPSKGPIMLGSVLSLRHSGMNOQVPLM 353
DB 241 GPATEAADLVTDGDEKQITLYGPBEGSGMIARPSKGPIMLGSVLSLRHSGMNOQVPLM 300
QY 354 LMDPIVTLFGSVHNNPQAGSGMRSTLFPNFGSMFVTDQAKNBOEMENLHRDDEYA 413
DB 301 LMDPIVTLFGSVHNNPQAGSGMRSTLFPNFGSMFVTDQAKNBOEMENLHRDDEYA 360
QY 414 SFGAGGDYEDNLHSPILSRQATGABGKDIVHGHGRSALSRRQTLLEGCGDVSTIDG 473
DB 361 SFGAGGDYEDNLHSPILSRQATGABGKDIVHGHGRSALSRRQTLLEGCGDVSTIDG 420
QY 474 GGMOLAMKSEKGENGRKGGPKRYLHQBGPVGRSGISYLSRGGGVFSGSEFVHA 533
DB 421 GGMOLAMKSEKGENGRKGGPKRYLHQBGPVGRSGISYLSRGGGVFSGSEFVHA 480
QY 534 ALVSGALFSGKGLAEPMSDAMVHPSEVAAGSRKMDLFEPCVRALLVGVGIQILOQF 593
DB 481 ALVSGALFSGKGLAEPMSDAMVHPSEVAAGSRKMDLFEPCVRALLVGVGIQILOQF 540
QY 594 AGINGVLYTTPQILEAGAVAVILSKFGLSASASIISSLTLLMLPCIGFAMLMMDLSG 653
DB 541 AGINGVLYTTPQILEAGAVAVILSKFGLSASASIISSLTLLMLPCIGFAMLMMDLSG 600
QY 654 RREFLLGTPIILIASVILVNSNLDIGTLAHLASTVSVIYFCCFWNGFPIPIILCA 713
DB 601 RREFLLGTPIILIASVILVNSNLDIGTLAHLASTVSVIYFCCFWNGFPIPIILCA 660
QY 714 ELPFTRVRGLCIAICAFTEWIGDIIVTYSLPVMLNAGLAGVFSIYAVVCLISFVFLK 773
DB 661 ELPFTRVRGLCIAICAFTEWIGDIIVTYSLPVMLNAGLAGVFSIYAVVCLISFVFLK 720
QY 774 VPEIKGMPLEVTTEPPAVAKAKOAAA 800
DB 721 VPEIKGMPLEVTTEPPAVAKAKOAAA 747

```

RESULT 2
US-09-291-922-8
Sequence 8, Application US/09291922
Patent No. 6383776
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Hitz, Bill
APPLICANT: Kinney, Tony
APPLICANT: Tingey, Scott
TITLE OF INVENTION: Plant Sugar Transport Proteins
FILE REFERENCE: BB-1163
CURRENT APPLICATION NUMBER: US/09/291,922
EARLIER FILING DATE: 1999-04-14
EARLIER APPLICATION NUMBER: 60/083,044
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Microsoft Office 97
SEQ ID NO 8
LENGTH: 737

```

; TYPE: PRT
; ORGANISM: Glycine max
US-09-291-922-8
Query Match 65.6%; Score 2674; DB 4; Length 737;
Best Local Similarity 69.9%; Pred. No. 8.9e+256;
Matches 522; Conservative 92; Mismatches 115; Indels 18; Gaps 8;

QY 56 MSGAVLVAIVASIGNLQGMNATTAANLYIKKEFOLONEPTVEGLIVSMGLIGATVIT 115
DB 1 MCGAVVVAIAASIGNLQGMNATTAAGAVLYIKKEFNLOSEPLIRGLIAAMFLIGATVIT 58
QY 116 TFSGLSDSIGRRPMLIISIIYFSGILMWSPNVYVLLARFVDPGFGIGLAATLVPLY 175
DB 59 TCSGFIADMLRRPMMIISVLYFLGGLVLMSPNVYVLLARFVDPGFGIGLAATLVPLY 118
QY 176 ISEIAPSEIRGILNLTLPQFSGSGGMFLSYCMVFGMSLSPSPDMRIMLGVALIPSLFFPGL 235
DB 119 ISEIAPSEIRSGINLTLPQFSGSGGMFLSYCMVFGMSLSPSPDMRIMLGVALIPSLFFPGL 178
QY 236 TIFPLPESPRWLVSKGRMAEAKKYLQRLRGKDVSGELSLLEGLVEGDTISIEBYI 295
DB 179 TIFPLPESPRWLVSKGRMAEAKKYLQRLRGKDVSGELSLLEGLVEGDTISIEBYI 238
QY 296 ATEAADLVTDGDEKQITLYGPBEGSGMIARPSKGPIMLGSVLSLRHSGMNOQVPLM 355
DB 239 ADVADGHEHATEKDKRLRYGSQAGLSWLSKPVVGQSSIG---LASHHGSIIINQSMPLM 294
QY 356 DPIVTLFGSVHNNPQAG--GSMRSTLFPNFGSMFVTDQAKNBOEMENLHRDDEYA 413
DB 295 DPLVTLFGSVHNNPQAG--GSMRSTLFPNFGSMFVTDQAKNBOEMENLHRDDEYA 354
QY 414 SDAGGDYEDNLHSPILSRQATGABGKDIVHGHGRSALSRRQTLLEGCGDVSTIDG 471
DB 355 SDATRGDSDNLHSPILSRQATGABGKDIVHGHGRSALSRRQTLLEGCGDVSTIDG 413
QY 472 ICGGMOLAMKSEKGENGRKGGPKRYLHQBGPVGRSGISYLSRGGGVFSGSEFVHA 531
DB 414 ICGGMOLAMKSEKGENGRKGGPKRYLHQBGPVGRSGISYLSRGGGVFSGSEFVHA 466
QY 532 AAALVSGALFSGKGLAEPMSDAMVHPSEVAAGSRKMDLFEPCVRALLVGVGIQILOQF 591
DB 467 AAALVSGALFSGKGLAEPMSDAMVHPSEVAAGSRKMDLFEPCVRALLVGVGIQILOQF 526
QY 592 QPAGINGVLYTTPQILEAGAVAVILSKFGLSASASIISSLTLLMLPCIGFAMLMMDL 651
DB 527 QPAGINGVLYTTPQILEAGAVAVILSKFGLSASASIISSLTLLMLPCIGFAMLMMDL 586
QY 652 SGRRLFLGTPIILIASVILVNSNLDIGTLAHLASTVSVIYFCCFWNGFPIPIILCA 711
DB 587 SGRRLFLGTPIILIASVILVNSNLDIGTLAHLASTVSVIYFCCFWNGFPIPIILCA 646
QY 712 CAEIFPTRVRGLCIAICAFTEWIGDIIVTYSLPVMLNAGLAGVFSIYAVVCLISFVFLK 771
DB 647 CAEIFPTRVRGLCIAICAFTEWIGDIIVTYSLPVMLNAGLAGVFSIYAVVCLISFVFLK 706
QY 772 LKVPETKMPLEVTTEPPAVAKAKOAAA 798
DB 707 LKVPETKMPLEVTTEPPAVAKAKOAAA 733

```

RESULT 3
US-09-291-922-29
Sequence 29, Application US/09291922
Patent No. 6383776
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Hitz, Bill
APPLICANT: Kinney, Tony
APPLICANT: Tingey, Scott
TITLE OF INVENTION: Plant Sugar Transport Proteins
FILE REFERENCE: BB-1163
CURRENT APPLICATION NUMBER: US/09/291,922

/ CURRENT FILING DATE: 1999-04-14
 / EARLIER APPLICATION NUMBER: 60/083,044
 / EARLIER FILING DATE: April 24, 1998
 / NUMBER OF SEQ ID NOS: 30
 / SOFTWARE: Microsoft Office 97
 / SEQ ID NO 29
 / LENGTH: 729
 / TYPE: PRT
 / ORGANISM: Arabidopsis thaliana
 US-09-291-922-29

Query Match 63.4%; Score 2584; DB 4; Length 729;
 Best Local Similarity 68.8%; Pred. No. 7,3e-247;
 Matches 512; Conservative 89; Mismatches 123; Indels 20; Gaps 8;

QY 56 MSGAVLVAVASIGNULQGMNATTAATLTKKEPOLONEPTVGLIYMSLIGATVIT 115
 DB 1 MSGAVLVAVAAVAGNULQGMNATTAAGAVLYIKKEFNLSNPSVEGLVAMSLIGATVIT 60
 QY 116 TFGCPILSDSIRGRRPMLISSILYFSGILMWSPNVYVLLARFVDFGIGLAVTLVPLY 175
 DB 61 TCGGAVADWLGRRPMLISSILYFVGSILVMSPNVYVLLARLDFGVLVTLVPLY 120
 QY 176 ISEIAPSEIRGLNLTLPQFSGSGMFLSYCMVFGMSLSPSPWRIMLGVLALPSLFFGL 235
 DB 121 ISEIAPSEIRGLNLTLPQFTGSGMFLSYCMVFGMSLSPSPWRIMLGVLALPSLFFGL 180
 QY 236 TIFPLPESPRMLVSGRMAEAKVLYOKLAKGKODVSESLLEGLVEGDTSEIETIYIP 295
 DB 181 TIFPLPESPRMLVSGRMAEAKVLYOKLAKGKODVSESLLEGLVEGDTSEIETIYIP 240
 QY 296 ATEAADLVTDGDEKQITLYGPEEGSWIARPSKGPIMLGVSILASRHGS-MVANSVPL 354
 DB 241 ABEVTDHDIAYDKQIKLYGABEGLSWARPVKG---GSTWVSLSTRGSTMRRQSL 296
 QY 355 MDPIYTLFGSVHNNPQAGSNRSTLFPNFGSMFVTDQAKNQEWDENLHRDDEEYAS 414
 DB 297 IDPLVTLFGSVHEKMPDT-GSMRSLFPHFGSMFVCGNQPHEDMDENLVGEGEDYS 355
 QY 415 DQAGDYEENHSPILSRATGAEKDIYHGHGSGALSMRQITLLEGGDVSTIDIG 474
 DB 356 D-HGDDSEDDHSPILSRQTSME-KDMPHTAHGLSTRHGSQVQAGGEGAGSMGIGG 413
 QY 475 GQOLAMKSEKGEKNGRKGFFRKYVYLHQGVGSGRGSIVLPGGADVFESEFYHAA 534
 DB 414 GQOLAMKSEKGEKNGRKGFFRKYVYLHQGVGSGRGSIVLPGGADVFESEFYHAA 462
 QY 535 LVSQSLFSGGLAEPAMDAAVHPSEVAAKGSRWKDLFEPGVRALLVGVGIQILQCPA 594
 DB 463 LVSQPALVSKDLKERTIGPAVHPSE-TTKGSIHMDLHPGVKALVVGVLQILQCPA 521
 QY 595 GINGVLYTPQILBQAGAVIISKGLSGASASILISLTLLMPCIFAMLMDSGR 654
 DB 522 GINGVLYTPQILBQAGAVIISKGLSGASASILISLTLLMPCIFAMLMDSGR 581
 QY 655 RFLLTGTPILASVILVSNLIDGTLAHLSTVSIVVFCCEVWGEPIPNILCAE 714
 DB 582 RFLLTGTPILASVILVSNLIDGTLAHLSTVSIVVFCCEVWGEPIPNILCAE 641
 QY 715 IFFTRVGLCIAFTFWIGDIIVTSLPVMALNAGLAVSIVAVCLISFVEVFLKY 774
 DB 642 IFFTRVGLCIAFTFWIGDIIVTSLPVMALNAGLAVSIVAVCLISFVEVFLKY 701
 QY 775 PETKMPLEVTTEFPVAGKAA 798
 DB 702 PETKMPLEVTTEFPVAGKAA 725

RESULT 4
 US-09-291-922-10
 / Sequence 10, Application US/09291922
 / Patent No. 6383776
 / GENERAL INFORMATION:

/ APPLICANT: Allen, Steve
 / APPLICANT: Hitz, Bill
 / APPLICANT: Kinney, Tony
 / APPLICANT: Kinney, Scott
 / TITLE OF INVENTION: Plant Sugar Transport Proteins
 / FILE REFERENCE: BB-1163
 / CURRENT APPLICATION NUMBER: US/09/291,922
 / CURRENT FILING DATE: 1999-04-14
 / EARLIER APPLICATION NUMBER: 60/083,044
 / EARLIER FILING DATE: April 24, 1998
 / NUMBER OF SEQ ID NOS: 30
 / SOFTWARE: Microsoft Office 97
 / SEQ ID NO 10
 / LENGTH: 486
 / TYPE: PRT
 / ORGANISM: Glycine max
 US-09-291-922-10

Query Match 41.0%; Score 1670.5; DB 4; Length 486;
 Best Local Similarity 68.1%; Pred. No. 1.4e-156;
 Matches 333; Conservative 57; Mismatches 84; Indels 15; Gaps 8;

QY 308 DKEQITLYGPEEGSWIARPSKGPIMLGVSILASRHGS-MVANSVPLMDPIVTLFGSVHE 367
 DB 5 EKDQITLYGPEEGSWIARPSKGPIMLGVSILASRHGS-MVANSVPLMDPIVTLFGSVHE 59
 QY 368 NMPQAGSNRSTLFPNFGSMFVTDQAKNQEWDENLHRDDEEYASDAGDVEDNLHS 427
 DB 60 KLPETG---STLFPHFGSMFVCGNQPHEDMDENLVGEGEDYS 113
 QY 428 PLTSROATGAEKDIYHGHGSGALSMRQITL-GEEDGVSTIDIGGQOLAMKSEK 486
 DB 114 PLTSROATGAEKDIYHGHGSGALSMRQITL-GEEDGVSTIDIGGQOLAMKSEK 171
 QY 487 GENGKKEGFRKYVYLHQGVGSGRGSIVLPGGADVFESEFYHAAVLSQSLFSGGL 546
 DB 112 GENGKKEGFRKYVYLHQGVGSGRGSIVLPGGADVFESEFYHAAVLSQSLFSGGL 230
 QY 547 AEPAMDAAVHPSEVAAKGSRWKDLFEPGVRALLVGVGIQILQCPA 606
 DB 221 AEPAMDAAVHPSEVAAKGSRWKDLFEPGVRALLVGVGIQILQCPA 290
 QY 607 LEQAGVGLSLSGISSTSSFLISAVTLLMPCIFAMLMDSGRFLITPILI 666
 DB 231 LEQAGVGLSLSGISSTSSFLISAVTLLMPCIFAMLMDSGRFLITPILI 350
 QY 351 AALTLVGLSLVDGSTANASISTISVIVFCFVWGEPIPNILCAEIPTRVGLCIA 410
 QY 727 ICAFTFWIGDIIVTSLPVMALNAGLAVSIVAVCLISFVEVFLKYPETKMPLEVT 786
 DB 411 ICAFTFWIGDIIVTSLPVMALNAGLAVSIVAVCLISFVEVFLKYPETKMPLEVT 470
 QY 787 EFPVAGK 795
 DB 471 EFPVAGK 479

RESULT 5
 US-09-291-922-14
 / Sequence 14, Application US/09291922
 / Patent No. 6383776
 / GENERAL INFORMATION:
 / APPLICANT: Allen, Steve
 / APPLICANT: Hitz, Bill
 / APPLICANT: Kinney, Tony
 / APPLICANT: Kinney, Scott
 / TITLE OF INVENTION: Plant Sugar Transport Proteins
 / FILE REFERENCE: BB-1163
 / CURRENT APPLICATION NUMBER: US/09/291,922
 / CURRENT FILING DATE: 1999-04-14
 / EARLIER APPLICATION NUMBER: 60/083,044

EARLIER FILING DATE: April 24, 1998
 NUMBER OF SEQ ID NOS: 30
 SOFTWARE: Microsoft Office 97
 SEQ ID NO 14
 LENGTH: 345
 TYPE: PR
 ORGANISM: Trifolium aestivum
 US-09-291-922-14

Query Match 34.9%; Score 1424; DB 4; Length 345;
 Best Local Similarity 82.9%; Pred. No. 2.1e-132;
 Matches 281; Conservative 24; Mismatches 32; Indels 2; Gaps 2;

QY 463 EGGDGVSSRDICGGQMLARKMSEKSEBNRKGKGFVVYVLRHOGVFGSRGSIYSLPGGG 521
 DB 4 EGGDAVSSGIGGGQMLARKMSEBROGDEKGFRRYVLRHOGVADSRGSIYSLPGGG 63
 QY 522 DVFE-GSEFVHAALVSSGALFSGKLAEPBM-SDAAMVHSEVAAGSRKDLFEFGVR 579
 DB 64 DATGGSGFTHAALVSHALYSKLMERMAAGPMITPLEAPKGSIMKDLFEFGVR 123
 QY 580 ALLVGVGIIQOFAGINCVLYTTPQILEQAGVAVILSKFGLSSASASILLSTLTML 639
 DB 124 ALFVGVGIIQOFAGINCVLYTTPQILEQAGVAVILSKFGLSSASASILLSTLTML 183
 QY 640 PCIGFAMLMDSGRRLLTGTPIILASIVLVSNLIDGLTAAALSTVSIVYFCC 699
 DB 184 PSIGVAMRMDISGRFLLGTPIILASIVLVSNLIDGLTAAALSTVSIVYFCC 243
 QY 700 FVMGFGPIINICAEIFPTVRGLCIACAFTFWDIIVTYSLPVMAIGLAGVFSIY 759
 DB 244 FVMGFGPIINICAEIFPTVRGVCAICATFTWICDIIIVTYSLPVMAIGLAGVFSIY 303
 QY 760 AVVCLISFVFLKVPETKGMPLVITEFFAVGAKQA 798
 DB 304 AVVCCIAFVFLKVPETKGMPLVITEFFAVGAKQA 342

RESULT 6
 US-09-291-922-16
 Sequence 16, Application US/09291922
 Patent No. 6383776
 GENERAL INFORMATION:
 APPLICANT: Allen, Steve
 APPLICANT: Hitz, Bill
 APPLICANT: Kinney, Tony
 APPLICANT: Tingey, Scott
 TITLE OF INVENTION: Plant Sugar Transport Proteins
 FILE REFERENCE: BB-1163
 CURRENT APPLICATION NUMBER: US/09/291,922
 CURRENT FILING DATE: 1999-04-14
 EARLIER APPLICATION NUMBER: 60/083,044
 EARLIER FILING DATE: April 24, 1998
 NUMBER OF SEQ ID NOS: 30
 SOFTWARE: Microsoft Office 97
 SEQ ID NO 16
 LENGTH: 228
 TYPE: PR
 ORGANISM: Trifolium aestivum
 US-09-291-922-16

Query Match 23.1%; Score 943; DB 4; Length 228;
 Best Local Similarity 80.6%; Pred. No. 5.1e-65;
 Matches 179; Conservative 25; Mismatches 18; Indels 0; Gaps 0;

QY 574 EPGVRAALVGVGIIQOFAGINCVLYTTPQILEQAGVAVILSKFGLSSASASILLSTLTML 633
 DB 1 EPGVRAALVGVGIIQOFAGINCVLYTTPQILEQAGVAVILSKFGLSSASASILLSTLTML 60
 QY 634 TTLMLPCIGFAMLMDSGRRLLTGTPIILASIVLVSNLIDGLTAAALSTVSIV 693
 DB 61 TTLMLPCIGFAMLMDSGRRLLTGTPIILASIVLVSNLIDGLTAAALSTVSIV 120

QY 694 IYVFCFVWGFGPIINICAEIFPTVRGLCIACAFTFWDIIVTYSLPVMAIGLAGVFSIY 753
 DB 121 IYVFCFVWGFGPIINICAEIFPTVRGLCIACAFTFWDIIVTYSLPVMAIGLAGVFSIY 180
 QY 754 GVFSIYAVVCLISFVFLKVPETKGMPLVITEFFAVGAKQA 795
 DB 181 GVFSIYAVVCLISFVFLKVPETKGMPLVITEFFAVGAKQA 222

RESULT 7
 US-09-291-922-6
 Sequence 6, Application US/09291922
 Patent No. 6383776
 GENERAL INFORMATION:
 APPLICANT: Allen, Steve
 APPLICANT: Hitz, Bill
 APPLICANT: Kinney, Tony
 APPLICANT: Tingey, Scott
 TITLE OF INVENTION: Plant Sugar Transport Proteins
 FILE REFERENCE: BB-1163
 CURRENT APPLICATION NUMBER: US/09/291,922
 CURRENT FILING DATE: 1999-04-14
 EARLIER APPLICATION NUMBER: 60/083,044
 EARLIER FILING DATE: April 24, 1998
 NUMBER OF SEQ ID NOS: 30
 SOFTWARE: Microsoft Office 97
 SEQ ID NO 6
 LENGTH: 131
 TYPE: PR
 ORGANISM: Oryza sativa
 US-09-291-922-6

Query Match 13.8%; Score 562; DB 4; Length 131;
 Best Local Similarity 80.5%; Pred. No. 1.2e-47;
 Matches 103; Conservative 17; Mismatches 8; Indels 0; Gaps 0;

QY 669 LVILVSNLIDGLTAAALSTVSIVYFCCFVMGFGPIINICAEIFPTVRGLCIAIC 728
 DB 1 LVILVSNLIDGLTAAALSTVSIVYFCCFVMGFGPIINICAEIFPTVRGLCIAIC 60
 QY 729 AFTFWIGDIIVTYSLPVMAIGLAGVFSIYAVVCLISFVFLKVPETKGMPLVITEF 788
 DB 61 AFTFWIGDIIVTYSLPVMAIGLAGVFSIYAVVCLISFVFLKVPETKGMPLVITEF 120
 QY 789 FAVGAKQA 796
 DB 121 FAVGAKQA 128

RESULT 8
 US-09-291-922-24
 Sequence 24, Application US/09291922
 Patent No. 6383776
 GENERAL INFORMATION:
 APPLICANT: Allen, Steve
 APPLICANT: Hitz, Bill
 APPLICANT: Kinney, Tony
 APPLICANT: Tingey, Scott
 TITLE OF INVENTION: Plant Sugar Transport Proteins
 FILE REFERENCE: BB-1163
 CURRENT APPLICATION NUMBER: US/09/291,922
 CURRENT FILING DATE: 1999-04-14
 EARLIER APPLICATION NUMBER: 60/083,044
 EARLIER FILING DATE: April 24, 1998
 NUMBER OF SEQ ID NOS: 30
 SOFTWARE: Microsoft Office 97
 SEQ ID NO 24
 LENGTH: 523
 TYPE: PR
 ORGANISM: Glycine max
 US-09-291-922-24

Query Match 12.2%; Score 496; DB 4; Length 523;

Best Local Similarity 21.9%; Pred. No. 4.7e-40;
Matches 165; Conservative 109; Mismatches 189; Indels 290; Gaps 18;

```

QY 59 AVVAIVASIGNLQGMNATIAAVALYIKKEFOLQNE--PTVELIYSMLIGATITTT 116
DB 31 AFACMLASMTSILGDIYGVSGAAYIKDKLSDQIEILLIGIIMLYSLIGSCL--- 87
QY 117 FSGPLSDSIGRRPMLISILYFSGGLIMLWSPNYVLLARFVDFGFGIGLAVTLVPLXI 176
DB 88 -AGRTSDMIGPRYITVFAGTIFPVGALLMGSPNPSFLMFGRFVAGIGIYALMAIAPYIT 146
QY 177 SEIAPSEIRGLNLTLPQFSSGGMFLSYCMVFGMS-LSPPSDWRIMGLVLAIPSLFFPGL 235
DB 147 AEVSPASRSGFLTSPFPEVINGGILIGIYSNVAFSKLTLYKGMRMMLGVGAIPSV-LTLTV 205
QY 236 TIFLIPESPRMLVSGRMAEAKVYLQKLRGKDDVSGELSLLEGLEVGDGDSIEEYITGP 295
DB 206 GVLMAPESPRLVMKGRLEAKVILNK----- 232
QY 296 ATEAADLVTDGDKQITLYGPEEGQSWIARPSKPIMLGSLVLSASHRGSMVNSVPLM 355
DB 233 -----TSDSKEE-----AQLRLA----- 245
QY 356 DPIVTLFGSVHNNPQAGSMRSTLFPNFGSMFVTDQHANQEWDEENLRHDEEYASD 415
DB 246 -----EIKQAAGIPESC-----NDD----- 260
QY 416 GAGGDYEDNLHSPILLSRQATGAEKDIYHHGHSALSMRQTLLEGGDGVSSDTIGGG 475
DB 261 -----VGVNKRQS-----NGEGV----- 273
QY 476 WQLAMKMEKEGENRKEGKGFRRVYLHQGVPSRSGISVSLPGGDVFEESBFVHAAL 535
DB 274 ----- 273
QY 536 VSQALFESKGLAEPMSDAAVHPSEVAAKSRWKDLF--EPGVRRALLVGVGIQLIQ 592
DB 274 -----WKELFLYPTPAIRHIVIALGHIHFQ 300
QY 593 FAGINGVLYTTPQILEQGVAVILSKFGLSASASILISLTLMLPCIGFAMLMDS 652
DB 301 ASGDAAVAVLYSPRIPEKAGIINDTKL-----LATVAVGFVYVFIILA-----ATFTLDV 351
QY 653 GRPFLIGTIPILASLYIVLVV--NLIDLG--TLAHLSTVS--YIVVFCGFVWGFGPI 708
DB 352 GRPFLISSVGGVSLTLTALISLTVIDHSEKLMWAGSSIAMALAVATFSIGAGIT 411
QY 709 NILCAEIPTRVGLCIAICAFTEWIGDIIVTYSLPVMLNAGLAGVFSIYAVCLISFV 768
DB 412 WYVSEIFPLRLRQGAAGAIVAVNRRTSAVVSMTLSLTRAITIGCAFVLCIATVGM 471
QY 769 FVFLKVPETKMPLEVIT-EFFVANGAKOAAAKA 800
DB 472 FFYTVLPETRGKLTLEDMEGSFGTFRSKSNASKA 504

```

RESULT 9
US-09-291-922-20

; Sequence 20, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Kinney, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; EARLIER FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97

; SEQ ID NO 20
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Zea mays
US-09-291-922-20

Query Match 12.0%; Score 489.5; DB 4; Length 513;
Best Local Similarity 21.6%; Pred. No. 2e-39;
Matches 169; Conservative 106; Mismatches 200; Indels 307; Gaps 17;

```

QY 46 VTSPDI---LEDKMSG---AVVAIVASIGNLQGMNATIAAVALYIKKEFOLQNEPT 98
DB 1 MASDELAKAVEPRKRGKXKASICALIASMASVILGDIYGVSGAAMAYIKKDLNI-TDVQ 59
QY 99 VEGILVMSLIGATITVTFSGP--LSDSIGRRPMLISILYFSGGLIMLWSPNYVLLA 157
DB 60 LEIILIGLISLY--SLFSPFAGARTSDIRGRLTVFAVIFPVGSLMGRAVNGMLMAG 117
QY 158 RFVDFGFGIGLAVTLVPLYISIAPESEIRGLNLTLPQFSSGGMFLSYCMVFGMSLSP-SP 216
DB 118 RFVAGVGVGGMIAIAPYTAIEISPAASRGFLTTPPEVFINIGILGILSNFAPARLPLHL 177
QY 217 DWRIMGLVLAIPSLFFFGTLTFYLPESPRMLVSGRMAEAKVYLQKLRGKDDVSGELSL 276
DB 178 GWRVMAIAGAVPS-GILALVFCMPESPRLVLKGRLDARAVALK----- 222
QY 277 LEGLEVGDGDSIEEYITIGPATEAADLVTDGDKQITLYGPEEGQSWIARPSKPIMLGS 336
DB 223 -----TSATPEEAERLA----- 235
QY 337 VLSASHRGSMVNSVPLMDPIVTLFGSVHNNPQAGSMRSTLFPNFGSMFVTDQHA 396
DB 236 ----- 235
QY 397 NEQWDEENLRHDEEYASDAGGDYEDNLHSPILLSRQATGAEKDIYHHGHSALSMRR 456
DB 236 ----- 235
QY 457 QTLLEGDDGVSSSTDIGGQWAMKMEKEGENRKEGKGFRRVYLHQGVPSRSGISV 516
DB 236 -----DIAAA-----AGIPKGLDGVT 253
QY 517 LP-----GGDVFEESBFVHAALVVSQALFESKGLAEPMSDAAVHPSEVAAKSRWKDL 572
DB 254 VPGEQGGGE-----QWKKL 270
QY 573 F--EPGVRRALLVGVGIQLIQFAGINGVLYTTPQILEQGVAVILSKFGLSASASITL 629
DB 271 ILSPTPAVRRLILSAVGLHFFQOAGSDSVQVYARLFKSAGITDDKKLGVTCA----- 325
QY 630 ISSLTLMLPCIGFAMLMDSGRFLIGTIPILASLYIL-----VSNLIDLGTLA 684
DB 326 VGVTKTFEIL-----VATFLLDRAGRRLILISTGCMIVSLICGSGLTVAGHHHDPTVAV 381
QY 685 HALSTVSIVYVFCGFVWGFGPINILCAEIPTRVGLCIAICAFTEWIGDIIVTYSLP 744
DB 382 AVALCIASTLSYIAFFSIGLGPITGVTSSEIFPLQVALGFAVGSANRRTSAVISMTFL 441
QY 745 VMLNAGIAGVFSIYAVVCLISFVFLKVPETKMPLEVITEF-----AVGAKQAA 797
DB 442 SLSKAITIGSFFLYSGIAAVAVWFFTCLEPTEGRLEBKGLFGMPDTGMAEADDA 501
QY 798 AK 799
DB 502 AK 503

```

RESULT 10

US-09-291-922-28
; Sequence 28, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve

```

; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 28
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Triticum aestivum
; US-09-291-922-28

Query Match      12.0%; Score 489; DB 4; Length 529;
Best Local Similarity 20.8%; Pred. No. 2,3e-39;
Matches 165; Conservative 106; Mismatches 215; Indels 306; Gaps 14;

QY 24 LPSVTLALPGPLPASCSSQEPVTSDDILEDKMSGAVLVAIVASIGNLQGMNATIAAA 83
DB 19 MASALPBEAGVHPRN-----KGNFKYAFCTALCASMATIVLGYDVGWMSGA 65
QY 84 VLYIKKEPQLONE--PVEGLIVMSLIGATVITTPSGPLSDSIGRRPMLLSILYEPFS 141
DB 66 SLYIKRDIQITDQLEIMMGILSVYALIGSFL---GARTSDWVRRTVYFAAAIFNNG 121
QY 142 GLIMMSRPVVYLLARFVDFGFGIGLAVTLVPLYISERISRGILNTLPOFGSGGMF 201
DB 122 SLIMFANVYALMNGRFVTGIGVYALMVAVITPEVSPAKRGFLISFTEVFIVNGIL 181
QY 202 LSYCNVFGKMSLSP-SPDWIRIMGLVLAIPSLFFFGLTITFLYLPESPMLVSKGMAAKVYL 260
DB 182 LGVSNVYAFARLPFLHSRKMVLMGIAVPSA-LIALNVFQMPESPRLVMKGRSLADARAVL 240
QY 261 OKLRKDVSGSLSLBESLEVGSDTSEIYIIGPATEADLVTDGDKQOITLYXGPREG 320
DB 241 ----- 240
QY 321 QSMIRPSKPIMLGSVLISLRHGSWVQSVPLMDPIVTLFGSVHNNPQAGSGMRSTL 380
DB 241 ----- 240
QY 381 FPNFGSMFSVTDQAKNEQWDEENLHRDEEYASDAGDYEDNLHSLPSQATGAEGK 440
DB 241 -----AKTSDTPEEAVERLDQIKCA----- 260
QY 441 DIVHGHGHSALSMRQTLLEGSDGVSTDIGGQOLAMKMSKEGENGRKGGFKRYV 500
DB 261 ----- 260
QY 501 LHOEGVPSGRSGIVSLPGGDDVFESEFVHAALVSQALPSKGLAEPMSDAMVHPS 560
DB 261 ---AGIPRELDGDVVMRK-----TKGQEKQV----- 285
QY 561 EVAAGKSMKDLF---EPGVRRALLVGVIQIILQOFAGINGVLYTTPQILEQAVAILLS 617
DB 286 -----WKELIFSPPTAMRRIILALGIHFQOATSDSVLVSPPVFSAGITGDNH 337
QY 618 KFGUSSASASILISSLTLLMLPCIGFAMLMDSGRFLLGTTIPILIASIV-----IL 672
DB 338 LLG---ATCAMGWMTLFI---VATPOLDRVGRRLTLTSTAGMLACILIGTGLT 388
QY 673 VVSNLIDGLTALHALSTVSIVVFCFVMGFGPIPNLICARIFPTRVRCIATCAITP 732
DB 389 VVGRHPDAKVPRAIGLCIVSILAYVSPSIGLITSVTSEVFLKRALGALGISTCN 448
QY 733 WIGDIIVTSLPVMNAIIGLGVFSIYAVVCLISFVFLKVPETKMPLEVITEFF--- 789
DB 449 RVTSAAVSMSPLSKALITIGSFVLYAGIAIIGIIPFTFIPETRGFLPEIIGLFLFMT 508

```

```

QY 790 --AVGAQQAAMK 799
DB 509 DTAVEADQDTATK 520

RESULT 11
US-09-291-922-30
; Sequence 30, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Kinney, Tony
; APPLICANT: Hitz, Bill
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 30
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Beta vulgaris
; US-09-291-922-30

Query Match      12.0%; Score 488; DB 4; Length 549;
Best Local Similarity 21.3%; Pred. No. 3,1e-39;
Matches 164; Conservative 113; Mismatches 194; Indels 298; Gaps 19;

QY 41 SSGEPVTSDDILED-----KMSGAVLVAIVASIGNLQGMNATIAAVLYIKKEF 91
DB 10 SDPEPTTASKVIADPDLKPKKPKKFAFACATLASVTLGIDVMSGAIITLKEW 69
QY 92 QLONEPTEGLIVMSLIGATVITTPSGPLSDSIGRRPMLLSILYFSGILMLGSPNV 151
DB 70 HISSD--TQIGVGLIINLYCLFGSFAAGRTSDWIGRRYTVLAKATFVGALMGFATNY 127
QY 152 VYLLARFVDFGFGIGLAVTLVPLYISERISRGILNTLPOFGSGGMFLSYCMVFGMS 211
DB 128 AFLWVGRFVTGIGVYALMIAPIYTAIVSPASSSGFLTSPFVYINAGILGYISNLAFS 187
QY 212 LSPSP-DWRIMGLVLAIPSLFFFGLTITFLYLPESPRLVSKGRMAAKVYOKLRKDVDS 270
DB 188 SLFTHLSWRFMGLGIGALPISI-FLAIGVLAIMPESPRLVMQGRIGLDAKKVLNR----- 238
QY 271 GELSLLLEGLEVGSDTSEIYIIGPATEADLVTDGDKQOITLYGPEGQSMIARPSK 330
DB 239 -----ISD-----SPEAO----- 247
QY 331 PIMLGVLISLRHGSWVQSVPLMDPIVTLFGSVHNNPQAGSGMRSTLFPNFGSMFSV 390
DB 248 ----- 247
QY 391 TDQAKNEQWDEENLHRDEEYASDAGDYEDNLHSLPSQATGAEGQIYHGHGHS 450
DB 248 -----L 248
QY 451 ALSMRQTLLEGSDGVSTDIGGQOLAMKMSKEGENGRKGGFKRYVILHOGVPSGR 510
DB 249 RLSEIKQT-----AGTAA-- 261
QY 511 RGSIVSLPGGDDVFESEFVHAALVSQALPSKGLAEPMSDAMVHPSVAAGKSRWK 570
DB 262 -----ECDE-----DIYKTEKTKISGNVWK 283
QY 571 DLF---EPGVRRALLVGVIQIILQOFAGINGVLYTTPQILEQAVAILSKFGLSASAS 627
DB 284 ELTFENPRAVRAVIAIGIHFQOASGIDAVVLYSPRISQAGITVARQK-----LAT 338
QY 628 ILLSLLTLLMLPCIGFAMLMDSGRFLLGTTIPILIASIVILVVS-NLIDLG-----T 682

```

Db 339 VAVGVKTLFIL---VATFQLDKXGRBPLLTSGVMIAITLAMSFTVIDSHHKT 394
Qy 663 LAHALSTVSVIVFCFVGGPPIPNILCAEIPTRVR--GLCAIACFTWIDIIYT 740
Db 395 WAIALCTM-VCAVVASFISIGLPIWVSSVFPRLAOGTSMGV-AVNRVSGVLSI 452
Qy 741 YSLPMLNAIGLVFSIYAVVCLISFVFLKVPETKGMPLVITEF 789
Db 453 FFLP-LSHKITTCGAFLEFGIAIIMFPFLPFLPRTORTLENNHELF 500

RESULT 12

US-09-291-922-22
; Sequence 22, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; EARLIER FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 22
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (102)
US-09-291-922-22

Query Match 11.9%; Score 485; DB 4; Length 510;

Best Local Similarity 22.8%; Pred. No. 5.5e-39;
Matches 173; Conservative 98; Mismatches 187; Indels 300; Gaps 19;

Qy 59 AVLVAIVASIGNILQGMNATIAAVALYIKKEPOLONEPTVE--GLIVSMGLIGATVY 115
Db 21 APACILIASMTSILGVDIGVMSGASLIYIKDPNT-SDKRVLMGIATLNLISGFA-- 77
Qy 116 TFSGLSDSIGRPMILISILYFSGILMSPNVYVLLARFYDGFGLAVTLVPLY 175
Db 78 --AGRTSDWIGRRTIVFAVIFPAGXFLMGFAVYVAMLMFGRFVAGIGVYALMIPVY 135
Qy 176 ISEIAPSEIRGLNLTLPQSSGGMFLSCVFGKSLSP-SPDWRTIMGLVALIPSLFFPG 234
Db 136 TAEVSPAPARGFLTSPPEVFINFGILIGVSNVAFSRPLPLNGWIRIMIGIGAPSV-LTA 194
Qy 235 LTIIFLPESPRLVSKGRMAEAKVLOKRGKDVSGELSLLEGLVEVGDTSEIYYIG 294
Db 195 LMLVGMFSPRMLVWKGRADAKVLEKT-----SDT----- 226
Qy 295 PATEAADLVTDGKEQITLTYGPEEGQSWIARPSKPIMLGSLVSLASHRGSMVQSVPL 354
Db 227 -AEBAERL----- 234
Qy 355 MDPIYTLGSGVHNMPOAGSMRSTLFPNFGSMFSTVDQAHAKNQWDEINLRDDEEYAS 414
Db 235 ----- 234
Qy 415 DGAGGDYEDNLHSLPLSRQATGAEKDIYHHGRGASLSMRQTLLEGEGDVSSDTIGG 474
Db 235 -----ADIVA 239
Qy 475 GQWLAKMSKEGNGRKEGFRVYVYHOGVPGSRGSI VSLP--GGGDVFEGBEFVHA 532
Db 240 A-----AGIPEILDGDVTVTPKRGSGN----- 261

Qy 533 AALVQSALFSKGLAEPMSDAAMVHPSEVAAKGRMKDLF---BQVRALLVGVGIQI 589
Db 252 -----EKV-----WELISPTPMKRILISIGIHF 289
Qy 590 LQOFAGINGVLYYTPQILEQAGVAVILSKFGLSASASILISLTLLMPCIGPAMLM 649
Db 290 FQHALGHSVVFYSPVFKSPGLT--NDKHFLGTWPFGVTKRFLTL-----ATFPI 340
Qy 650 DLGSRPILIGTIPILASLV-----ILVSNLIDLGLAALASTVSVIYFCCFWMGF 704
Db 341 DQVGRPILIGTIGIILSLIGLAGLTVGQHPDAKIPMAIGLISASTLAVVAFSICL 400
Qy 705 GPIPIILCAEIPTRVR--GLCAIACFTWIDIIYVSLPYMNAIIGLVFSIYAVV 762
Db 401 GPITWVSEIFPLQVRLGSLGVANRVTSGVISMFTL--SLSKAITIGSFFLYSGI 458
Qy 763 CLISFVFLKVPETKGMPLVITEFPAVGAOKAOKA 800
Db 459 AALAVFFYTLPLPRTORTLEMSKLF--GD7AAASES 494

RESULT 13

US-09-291-922-26
; Sequence 26, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; EARLIER FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 26
; LENGTH: 539
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-291-922-26

Query Match 11.4%; Score 463.5; DB 4; Length 539;

Best Local Similarity 21.7%; Pred. No. 8.2e-37;
Matches 166; Conservative 108; Mismatches 200; Indels 291; Gaps 17;

Qy 52 LEDKMSG---AVLVAIVASIGNILQGMNATIAAVALYIKKEPOLONEPTVE--GLIV 104
Db 35 VEPKKGVVRPAFCALIASMTSILGVDIGVMSGASLIYIKDKLT-NDTQLEVLGILN 93
Qy 105 SMSLIGATVITTFSGPLSDSIGRPMILISILYFSGILMSPNVYVLLARFYDGF 164
Db 94 VYSLGSAFA---AGRTSDWIGRRTIVFAVIFPAGLIMGFSVNYAMLMFGRVAGIG 149
Qy 165 IGLAVTLVLYISLSEIRGLNLTLPQSSGGMFLSCVFGWS-LSPSPDWRIWLG 223
Db 150 VGVALMIAPVNTGVEVSPASARGVLTSPPEVFINFGILIGVSNFAFARLSRLGWRIMLG 209
Qy 224 VLAIPLFFPGITFYLPESPRLVSKGRMAEAKVLOKRGKDVSGELSLLEGLBVG 283
Db 210 IGAIVSV-LIAPMVYGMFSPRMLVWKGRADAKVLAKT----- 248
Qy 284 GDTSEIYYIIGPATEAADLVTDGKEQITLTYGPEEGQSWIARPSKPIMLGSLVSLASR 343
Db 249 SDT-----PEBAERLA----- 260
Qy 344 HGSVMQSVPLMDPIVTLFGSVHNMPOAGSMRSTLFPNFGSMFSTVDQAHAKNQWDEE 403
Db 261 ----- 260

QY 404 NIHRDEEYASDAGDYEDNLSPLSRQATGAEGKDIVHGHGSLMSRKQTLLEG 463
 Db 261 -----DITKAA----- 266
 QY 464 GDGVSTDIGGQWOLAMKWEKEGNGRKEGFKRVYLHQEGVPSGRSIVSLPGGDPV 523
 Db 267 -----GIPGLDQDDVPVK----- 281
 QY 524 FEGSEFVHAALVSQALFSKGLAEPRMSDAAVHSEVAAGKSRWDLFEFGVRAALIV 583
 Db 282 -----NKGSSSEKRYLKDILSPITIA-----MHILIA 309
 QY 584 GVGIGIILQOFAGINGVLYTTPQILEAGAVAVLSKRGSS-ASASILISLTLMLPRT 642
 Db 310 GGIHFPGQSSGIDAVVLSPLVFKSAGIT-----GDSRLRGTTVAAGATVTFIL--- 360
 QY 643 GFAMLLMDLSGRFFLLGTIPILIASLV-----ILVSNLIDLGTALHALSTVSVIYVF 697
 Db 361 -VATFLDRIRRRPLVLTSTGMLVSLVGLATGLTIVISHHPDEKITWALVLCIFCIMAVY 419
 QY 698 CCFVNGFGEIPNILCAEIFPTRVRCGLCAICAPTFWIGDIIVTYSLPVMLNAILAGVPS 757
 Db 420 AFPSIGLGITWVYSEIFPLHVRALGCSIGVAVNALTSGVISMTFISLSKMTIGGAF 479
 QY 758 IYAVVCLISFVFLKVPETKMPLEVITEFPAVGA--KQAAKA 800
 Db 480 LFAGIASFAWVFVFLPETRGRTLEDMSLFGNTATYHKGAAEA 524

RESULT 14

US-08-928-692-13
 ; Sequence 13, Application US/08928692
 ; Patent No. 5958727

GENERAL INFORMATION:

APPLICANT: Brody, Howard
 APPLICANT: Yaver, Deborah S.
 APPLICANT: Lamsa, Michael
 APPLICANT: Hansen, Kim
 TITLE OF INVENTION: Methode for Modifying the Production of
 NUMBER OF SEQUENCES: 80
 CORRESPONDENCE ADDRESS:
 ADDRESS: No. 59587270 No. 5958727diak o No. 5958727th America, Inc.
 STREET: 405 Lexington Avenue
 CITY: New York
 STATE: NY
 COUNTRY: USA
 ZIP: 10174

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/928,692
 FILING DATE: 12-SEPT-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 NAME: Lambirth, Elias J
 REGISTRATION NUMBER: 33,728
 REFERENCE/DOCKET NUMBER: 4944.200-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-867-0123
 TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:
 LENGTH: 584 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: No. 5958727e
 US-08-928-692-13

Query Match 10.3%; Score 421.5; DB 2; Length 584;
 Best Local Similarity 19.1%; Pred. No. 1,4e-32;
 Matches 150; Conservative 121; Mismatches 195; Indels 321; Gaps 19;

QY 44 EPTVSD-----ILEDGMSGAVL-VATIASIGNLQGMNDATIAAVALYKKEFOLONE 96
 Db 64 KPVNDEDDTSVMITFNQSLSPFITITLTFVABISGFMGYDGYISALISIGTD--LDHK 121
 QY 97 PTVBG--LIVSMILGATVITTSGLSDSISGRPMILISIIYFSGILMNSPNVY 153
 Db 122 VLTGKEKIVTAATSLGALITSIPAGTAADIFGRKRCIMSGNSLMFVIGALIIQVSAHFQW 181
 QY 154 LLARFVDSGICGLAVTLVPLXISEIASEIRGLNTLTPQSSGCMFLSCYMPGMSLS 213
 Db 182 MAVGRLINGFVGIGSLAPLFIISLAPKMRGLTYNSIMLTGGLVAGCGAGLNY- 240
 QY 214 PSPMRIMLVIAIPSLFFPGTLTFYLPESPRMVLVSKRMAAKVLOKRGKDVSGEL 273
 Db 241 VNMGRILVGLSLIPVNOF--TCLCFPLDPTFRYYVMGDLARATEVLK--RSYDTSEI 297
 QY 274 SLLEGLEVGDSIIEYIIGPATBADDLVTDGKEQITLYGPBGQSMIARPSKPFIM 333
 Db 298 -----IERKV-----BELVT----- 307
 QY 334 LGSVLASLRGSMVNOGVPLMDPIVTLFGSVHNMPOAGSGMRSTLEPNFGSMFSVTDQ 393
 Db 308 -----LNOSIP----- 313
 QY 394 HAKOEQWDEENLHRDDEEYASDAGDYEDNLSPLSRQATGAEGKDIVHGHGSLMS 453
 Db 314 -----GNV----- 317
 QY 454 MRQTLLEGDGVSSTDIGGQWOLAMKWEKEGNGRKEGFKRVYLHQEGVPSGRGS 513
 Db 318 ----- 317
 QY 514 IVSLPGGSDVEGSEFVHAALVSQALFSKGLAEPRMSDAAVHSEVAAGKSRWDLF 573
 Db 318 -----PEKV-----WNTIK 326
 QY 574 E-----PGVRALVGVGIQLIQFAGINGVLYTTPQILEAGAVAVLSKFGSSASASIL 629
 Db 327 ELHVPSVSLRALIIGCGIAIQOFTGMSLSMYFSGITFETVYGFK-----NSAVSII 378
 QY 630 ISSLTLLMLPCIGFAMLLMDLSGRFFLLGTIPILIASLVIL-----VSN 676
 Db 379 VSGTNFIPTL-----VAFSIDKIGRRTILIGLPGMTALVCSIAFHFLIKEDGAVAV 434
 QY 677 LIDGTLAHLALSTVSVVYFCFVMGGRPIPNILCAEIFPTRVRCGLCAICAPTFWIGD 736
 Db 435 VSSGFSWGIIVITVFAFALGIVP-WQOSELFPQNVRGISTSYATATWAGS 493
 QY 737 IIVTSLPVMNAILGAVFSIYAVVCLISFVFLKVPETKMPLE-----VITEFPAV 792
 Db 494 LVIASTFLTMQNTIPACTFAFPAGLSCTIIFCYFCPELSGLEBEVQILMDGFIK 553

RESULT 15

US-09-339-972-13
 ; Sequence 13, Application US/09339972
 ; Patent No. 6323002

GENERAL INFORMATION:

APPLICANT: Brody, Howard
 APPLICANT: Yaver, Deborah S.
 APPLICANT: Lamsa, Michael
 APPLICANT: Hansen, Kim
 TITLE OF INVENTION: Methode for Modifying the Production of
 NUMBER OF SEQUENCES: 80

CORRESPONDENCE ADDRESS:
ADDRESSES: No. 63230020 No. 6323002diak of No. 6323002ch America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/339,972
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/928,692
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4944,200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 584 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6323002e
US-09-339-972-13

Query Match 10.3%; Score 421.5; DB 4; Length 584;
Best Local Similarity 19.1%; Pred. No. 1.4e-32;
Matches 150; Conservative 121; Mismatches 195; Indels 321; Gaps 19;

QY 44 ERYTDD-----IIEKMSGAVL-VAIVASIGNLQGMNDNTAAVLYIKKEPOLONE 96
DB 64 KVNDEDDTSMITRNQSLSPRIITLTFVASISGFMFGYDTGVISSALISIGTD--LDHK 121
QY 97 PTVEG---LIVMSLIGATVITTFSGPLSDSIGRRPMLISILVFFSGILMLMSGPNVYV 153
DB 122 VLTVEKEIVTATSLGALITISIFACTADIFGRKCLMGLMVFYIGAILVOVSAHTWQ 181
QY 154 LILAFVDFGIGLAVTLVPLVYISRIAPSEIRGLNLTLPQFSGSGMPLSYCMVFQMSLS 213
DB 182 MAVGRILMGFGVIGISLAPFLISRIAPMIRGLTVINSMLTGGQLVAYCGAGLNY- 240
QY 214 PSPDRIMLGVALPSLFFGLITTYLPSPRMLVSKGMMAKAYLOKLRKGDVSGEL 273
DB 241 VNNGRRIIVGLSLIPAVQF-TCLCFLPDTPRYVWKGDLPARATEVLK--RSYTDTSERI 297
QY 274 SLLLEGLEVGDTSEEBYIGPATEADLVNDGXEQITLGPREGSGWIARPSKPIIM 333
DB 298 -----IERKV-----BELVT----- 307
QY 334 LGSVLISLASRHGSMVNOVPLMDPIVTLFGSVHNNPQAGSGMRSTLPNPGSMFVSVDQ 393
DB 308 -----LNQSIIP----- 313
QY 394 HAKNEQMBENLHRDDEYASDAGGDYEDNLHSPILSRQATGARGKDIIVHGHRSALS 453
DB 314 -----GKNV----- 317
QY 454 NKRQTLLEGGDGVASTDIGGMQLAMKWSKEGENGRKEGGFKRYVYLHQEGVPSRRGS 513
DB 318 ----- 317
QY 514 IVSLPGGDVFEGBEFVHAALVSGALPFSKGLAEPRMSDAAVHPSEVAAGSRWKDLF 573

DB 318 -----PERV-----WNTIX 326
QY 574 E-----PGVRALIVGVGIIQLQCPAGINGVLYYTPQILQAGVAVILSKFGLSSASATL 629
DB 327 ELHTVPSNLRALITIGCGLOAIQOFTGMNSLMTYFSGTIFETVGRK-----NSSAVSII 378
QY 630 ISSLTTLMLPCIGFAMLMNDLSGRFPILGTIPILIASVIL-----VVSN 676
DB 379 VSGTNFIETL-----VAFPSIDKIGRRTILLGLGGMALVCSIAHFLGIKFDGAVAV 434
QY 677 LIDGTLAHLSTVSIVYVFCFWMGFRPIINILCAIFPTRVRLCIAICATFWIGD 736
DB 435 VSSGFSWGIIVIFVIFVAFALGIVP-WQOSELPQNVRGIGTSYATATNAGS 493
QY 737 IIVYTSLEVMNAIIGLAVPSIYAVVCLISPFVFLKVPETKGMPL-----VITEFPAYG 792
DB 494 LVIASTFLTMQNTIPAGTFAFFAGLSCLSTIFCYFCELSGLELBEVQITLKDGRNIK 553
QY 793 AKQAAK 799
DB 554 ASKALAK 560

Search completed: January 5, 2004, 18:59:28
Job time : 37 secs

THIS PAGE BLANK (058710)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2004, 14:13:32 : Search time 867 Seconds
(without alignments)
11086.463 Million cell updates/sec

Title: US-10-051-909-31

Perfect score: 2777
Sequence: 1 caccgggttagctcgagc.....aaaaaaaaaaaaaaaaaa 2777

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2263443 seqs, 1730637950 residues

Total number of hits satisfying chosen parameters: 4526866

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the local score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2777	100.0	2777	14	US-10-051-909-31
2	2740.4	98.7	2908	12	US-10-310-154-350
3	2200.2	79.2	2824	14	US-10-051-902-1
4	2200.2	79.2	2824	14	US-10-051-909-1
5	940.2	33.9	2601	14	US-10-051-902-7
6	940.2	33.9	2601	14	US-10-051-909-7
7	875.8	31.5	2190	10	US-09-938-842A-1315
8	700	25.2	2205	10	US-09-938-842A-2254
9	691.2	24.9	1487	14	US-10-051-902-13
10	691.2	24.9	1487	14	US-10-051-909-13
11	594.6	21.4	1682	14	US-10-051-902-9
12	594.6	21.4	1682	14	US-10-051-909-9
13	394.4	14.2	1009	14	US-10-051-902-15
14	394.4	14.2	1009	14	US-10-051-909-15
15	269.2	9.7	285	9	US-09-923-876-6145

16	269.2	9.7	285	12	US-09-923-876-6145	Sequence 6145, Ap
17	257.2	9.3	282	9	US-09-294-093B-699	Sequence 699, App
18	227	8.2	870	14	US-10-051-902-5	Sequence 5, Appl
19	227	8.2	870	14	US-10-051-909-5	Sequence 5, Appl
20	123.2	4.4	510	14	US-10-051-902-11	Sequence 11, Appl
21	123.2	4.4	510	14	US-10-051-909-11	Sequence 11, Appl
22	114.2	4.1	443	14	US-10-051-902-3	Sequence 3, Appl
23	114.2	4.1	443	14	US-10-051-909-3	Sequence 3, Appl
24	112.4	4.0	299	9	US-09-294-093B-6094	Sequence 6094, Ap
25	96.2	3.5	1853	14	US-10-051-902-23	Sequence 23, Appl
26	96.2	3.5	1853	14	US-10-051-909-23	Sequence 23, Appl
27	84	3.0	2017	14	US-10-051-902-21	Sequence 21, Appl
28	84	3.0	2017	14	US-10-051-909-21	Sequence 21, Appl
29	82.6	3.0	2051	12	US-10-310-154-356	Sequence 356, App
30	81.4	2.9	260	10	US-09-878-574-9951	Sequence 9951, Ap
31	78.2	2.8	2089	14	US-10-051-902-25	Sequence 25, Appl
32	78.2	2.8	2089	14	US-10-051-909-25	Sequence 25, Appl
33	77.6	2.8	1482	10	US-09-938-842A-795	Sequence 795, App
34	75	2.7	1644	10	US-09-938-842A-320	Sequence 320, App
35	66.2	2.4	1386	12	US-10-369-493-47058	Sequence 47058, A
36	65.8	2.4	1872	14	US-10-051-902-27	Sequence 27, Appl
37	65.8	2.4	1872	14	US-10-051-909-27	Sequence 27, Appl
38	65.2	2.3	1655	12	US-10-369-493-26996	Sequence 26996, A
39	64.2	2.3	1422	12	US-10-369-493-46784	Sequence 46784, A
40	63.4	2.3	1377	12	US-10-369-493-39348	Sequence 39348, A
41	63.4	2.3	1377	12	US-10-369-493-39739	Sequence 39739, A
42	63.4	2.3	1419	12	US-10-369-493-38982	Sequence 38982, A
43	63.4	2.3	1914	14	US-10-051-902-19	Sequence 19, Appl
44	63.4	2.3	1914	14	US-10-051-909-19	Sequence 19, Appl
45	63.4	2.3	2127	12	US-10-310-154-354	Sequence 354, App

ALIGNMENTS

RESULT 1

US-10-051-909-31
Sequence 31, Application US/10051909
Publication No. US20020199217A1
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Helentjaris, Tim
APPLICANT: Hitz, Bill
APPLICANT: Kinney, Tony
APPLICANT: Tingey, Scott
TITLE OF INVENTION: Plant Sugar Transport Proteins
FILE REFERENCE: B01163 US CIP
CURRENT APPLICATION NUMBER: US/10/051,909
CURRENT FILING DATE: 2002-01-17
PRIOR APPLICATION NUMBER: 60/083,044
PRIOR FILING DATE: April 24, 1998
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Microsoft Office 97
SEQ ID NO 31
LENGTH: 2777
TYPE: DNA
ORGANISM: Zea mays
US-10-051-909-31

Query Match 100.0%: Score 2777, DB 14, Length 2777;
Best Local Similarity 100.0%: Pred. No. 0;
Matches 2777; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACGGGATTAGATTGAGAGCGGCTTGTGACGTCGCAACGCCCTTACCCCTGATC 60
DB 1 CACGGGATTAGATTGAGAGCGGCTTGTGACGTCGCAACGCCCTTACCCCTGATC 60
QY 61 TGGACCGGAGGAGGCGCTTCCGTCAGTGTCTTGTGCTGGCGCTTCCGCTTCCG 120
DB 61 TGGACCGGAGGAGGCGCTTCCGTCAGTGTCTTGTGCTGGCGCTTCCGCTTCCG 120
QY 121 CTGCTTCTGTTCTTCTTACAGAGCGGCTGACCTCGGACGATATCTTGGAGGACAAGATGT 180
DB 121 CTGCTTCTGTTCTTCTTACAGAGCGGCTGACCTCGGACGATATCTTGGAGGACAAGATGT 180

Db 121 CTGCTTCGTTCTTCACAGAGCCGGTGAACCTCGACGATATCTTGAGAGACAAGATGT 180
Qy 181 CGGGGGCTGTTCTTGCCGATAGTCGCTCCATCGGCAATCTATTGCAAGGGTGGACA 240
Db 181 CGGGGGCTGTTCTTGCCGATAGTCGCTCCATCGGCAATCTATTGCAAGGGTGGACA 240
Qy 241 ATGGCAACCATCCAGCTGCTGTTCTGATATATAAGAAAGAAATTTCAATTTGCAAAATGAGC 300
Db 241 ATGGCAACCATCCAGCTGCTGTTCTGATATATAAGAAAGAAATTTCAATTTGCAAAATGAGC 300
Qy 301 CCACTGTGAGGAGCACTAATTTGTCAATGTCACTTATCGGGCCACCATCGTTACTACAT 360
Db 301 CCACTGTGAGGAGCACTAATTTGTCAATGTCACTTATCGGGCCACCATCGTTACTACAT 360
Qy 361 TCTCCGGGCCATTATACAGACTGATTTGGCCGACGCCCTATGCTTATTTCTCTTCAATTC 420
Db 361 TCTCCGGGCCATTATACAGACTGATTTGGCCGACGCCCTATGCTTATTTCTCTTCAATTC 420
Qy 421 TGTACTCTTTCAGCGGCTCATGATGCTATGCTCTTAATGTCTAATGTCTAATGTCTGCTGTGG 480
Db 421 TGTACTCTTTCAGCGGCTCATGATGCTATGCTCTTAATGTCTAATGTCTAATGTCTGCTGTGG 480
Qy 481 CAGCTTCGTAGATGATTTGTGATTTGGCTGTCAAGCTTGTGCTTGTACATTT 540
Db 481 CAGCTTCGTAGATGATTTGTGATTTGGCTGTCAAGCTTGTGCTTGTACATTT 540
Qy 541 CAGAAATAGCCCCCTTCGAGATTTAGAGTTTGCTGAATACTACCACAATTCAGTGAAT 600
Db 541 CAGAAATAGCCCCCTTCGAGATTTAGAGTTTGCTGAATACTACCACAATTCAGTGAAT 600
Qy 541 CAGAAATAGCCCCCTTCGAGATTTAGAGTTTGCTGAATACTACCACAATTCAGTGAAT 600
Db 541 CAGAAATAGCCCCCTTCGAGATTTAGAGTTTGCTGAATACTACCACAATTCAGTGAAT 600
Qy 601 CAGGAGGAATGTTCTGTGATCTGATAGTGTGTTGGAGATGTCCTGTGGCAATCCCG 660
Db 601 CAGGAGGAATGTTCTGTGATCTGATAGTGTGTTGGAGATGTCCTGTGGCAATCCCG 660
Qy 661 ATTGAGATTAATGCTGTGCTGCTGCGATACCTTATTTCTTCTTGGTTGACAA 720
Db 661 ATTGAGATTAATGCTGTGCTGCTGCGATACCTTATTTCTTCTTGGTTGACAA 720
Qy 721 TATTTTATCTTCTGATCTCCAAAGATGCTGTTAGCAAAAGTGGATGGACAGGCA 780
Db 721 TATTTTATCTTCTGATCTCCAAAGATGCTGTTAGCAAAAGTGGATGGACAGGCA 780
Qy 781 AAAAGGTGTGCAAAAGTTACGGGGGGAAGAAGATCTCAGGTGAATGTCCTTCTTC 840
Db 781 AAAAGGTGTGCAAAAGTTACGGGGGGAAGAAGATCTCAGGTGAATGTCCTTCTTC 840
Qy 841 TCGAAGGTTGGAGTTGGAGAGACACTTCCATTGAGAGTACATCATTTGACCTGCA 900
Db 841 TCGAAGGTTGGAGTTGGAGAGACACTTCCATTGAGAGTACATCATTTGACCTGCA 900
Qy 901 CCGAGGAGCCCATGATCTTTGTTACTGACGGTGTATAGAAACAAATCACTTTATGGC 960
Db 901 CCGAGGAGCCCATGATCTTTGTTACTGACGGTGTATAGAAACAAATCACTTTATGGC 960
Qy 961 CTGAAGAGCCAGTCATGATTTGCTCGACCTTCTTAAGGAGCCCATCATGCTTGAAGTG 1020
Db 961 CTGAAGAGCCAGTCATGATTTGCTCGACCTTCTTAAGGAGCCCATCATGCTTGAAGTG 1020
Qy 1021 TGTCTTCTCTGATCTGTCATGAGAGATGTTGACCAAGATGTAACCCCTTATGAGATC 1080
Db 1021 TGTCTTCTCTGATCTGTCATGAGAGATGTTGACCAAGATGTAACCCCTTATGAGATC 1080
Qy 1081 CGATTTGTGACACTTTTGTGATGTCCATGAGAAATATGCTCAAGCTGAGAGAAATGTA 1140
Db 1081 CGATTTGTGACACTTTTGTGATGTCCATGAGAAATATGCTCAAGCTGAGAGAAATGTA 1140
Qy 1141 GAGAGCAATTTTCCAACTTTGGAAGTATGTCAGTGTCAAGATCAGATGCCAAA 1200
Db 1141 GAGAGCAATTTTCCAACTTTGGAAGTATGTCAGTGTCAAGATCAGATGCCAAA 1200
Qy 1201 ATGAGCAGTGGAGTGAAGAAATCTTCATAGGATACAGAGAGTACGATCTGATGTG 1260
Db 1201 ATGAGCAGTGGAGTGAAGAAATCTTCATAGGATACAGAGAGTACGATCTGATGTG 1260

Qy 1261 CAGAGGTGACTATATAGAGCAATCTCCATAGCCCATTTGCTGTCCAGGAGCAAGGTG 1320
Db 1261 CAGAGGTGACTATATAGAGCAATCTCCATAGCCCATTTGCTGTCCAGGAGCAAGGTG 1320
Qy 1321 CGAAGGAGAGGACATTGTGACCATTTGTCAACCGTGAAGTCTTGTAGCATGAGAAAGC 1380
Db 1321 CGAAGGAGAGGACATTGTGACCATTTGTCAACCGTGAAGTCTTGTAGCATGAGAAAGC 1380
Qy 1381 AAAACCTTTAGGGAGGGTGGAGATGTTGTGAGCAGACTGATATCGGTGGGATGAC 1440
Db 1381 AAAACCTTTAGGGAGGGTGGAGATGTTGTGAGCAGACTGATATCGGTGGGATGAC 1440
Qy 1441 AGCTTGTGGAATATGTCAGAGAGAGAGAGTGAATATGATAGAAAGAGTGTTC 1500
Db 1441 AGCTTGTGGAATATGTCAGAGAGAGAGAGTGAATATGATAGAAAGAGTGTTC 1500
Qy 1501 AAAAGTCTACTTGCACCAAGAGGAGTCTGAGCTCAAGAAAGGGGCTCAATTTGTTAC 1560
Db 1501 AAAAGTCTACTTGCACCAAGAGGAGTCTGAGCTCAAGAAAGGGGCTCAATTTGTTAC 1560
Qy 1561 TTCCGGTGTGCGCATTTTTTGAAGGTAGTGAATTTGTACATGCTGCTGTTAGTAA 1620
Db 1561 TTCCGGTGTGCGCATTTTTTGAAGGTAGTGAATTTGTACATGCTGCTGTTAGTAA 1620
Qy 1621 GTCACTGACACTTTTCTCAAAAGGTCCTTGTGAACCAAGCATGTCAATGTGCAATGG 1680
Db 1621 GTCACTGACACTTTTCTCAAAAGGTCCTTGTGAACCAAGCATGTCAATGTGCAATGG 1680
Qy 1681 TTCAACCATCGAGGTAGCTGCAAAAGGTTACCGTTGAAAGATTTGTTGAACCTGAG 1740
Db 1681 TTCAACCATCGAGGTAGCTGCAAAAGGTTACCGTTGAAAGATTTGTTGAACCTGAG 1740
Qy 1741 TGAGGCGTCCCTGTTAGTCTGCTGTGGAATTCAGATCTTCAACAGTTTGTGGAATAA 1800
Db 1741 TGAGGCGTCCCTGTTAGTCTGCTGTGGAATTCAGATCTTCAACAGTTTGTGGAATAA 1800
Qy 1801 ACGGTGTTCTGATCTATACCCCAAAATTTCTTGAAGCAAGCTGTGTGCAATTTCTTT 1860
Db 1801 ACGGTGTTCTGATCTATACCCCAAAATTTCTTGAAGCAAGCTGTGTGCAATTTCTTT 1860
Qy 1861 CCAATTTTGTCTGAGCTCGGATCAGCATCATCTTATGATCAGTTCTCTCACTACCTTAC 1920
Db 1861 CCAATTTTGTCTGAGCTCGGATCAGCATCATCTTATGATCAGTTCTCTCACTACCTTAC 1920
Qy 1921 TAAATCTTCTTGTGATGCTTTGTGCAATGCTGTTAGATCTTTCGGAAGAGTCTTT 1980
Db 1921 TAAATCTTCTTGTGATGCTTTGTGCAATGCTGTTAGATCTTTCGGAAGAGTCTTT 1980
Qy 1981 TGTGCTAGGACCAATTCATCTTGAATAGCATCTTATGATATCTGCTGTGTCCATTC 2040
Db 1981 TGTGCTAGGACCAATTCATCTTGAATAGCATCTTATGATATCTGCTGTGTGTCCATTC 2040
Qy 2041 TAAATGATTTGGGTGACATAGCCCATGCTTGTCTCCACCGTCAAGTGTATGCTACT 2100
Db 2041 TAAATGATTTGGGTGACATAGCCCATGCTTGTCTCCACCGTCAAGTGTATGCTACT 2100
Qy 2101 TCTGTGCTTGTGATAGGAAATTTGTGTCATTCCTCAATTTATGTGCAAGATCTTTC 2160
Db 2101 TCTGTGCTTGTGATAGGAAATTTGTGTCATTCCTCAATTTATGTGCAAGATCTTTC 2160
Qy 2161 CAACGAGGTCGTGCTGCTGTATGTCATTTGTGCTTTAATCTGTGATCGGAGTAA 2220
Db 2161 CAACGAGGTCGTGCTGCTGTATGTCATTTGTGCTTTAATCTGTGATCGGAGTAA 2220
Qy 2221 TCAATGACCTACAGCTTCTGTGATGCTGAATGATTTGAGTGTGAGCTGCGGGTGTTC 2280
Db 2221 TCAATGACCTACAGCTTCTGTGATGCTGAATGATTTGAGTGTGAGCTGCGGGTGTTC 2280
Qy 2281 GCATATATGAGTGTATGCTTATTTCTTGTGTTGCTTCTTAAAGTCTCTGAGA 2340
Db 2281 GCATATATGAGTGTATGCTTATTTCTTGTGTTGCTTCTTAAAGTCTCTGAGA 2340

Db 759 AGGAATGTTCTTGTCTACTGATGATGTTGGGATGTCCTGTCGCCATCAACCCGATG 818
Qy 665 GAGAAATTAATGCTGTTGCTGCTGCGATACCTTCATGTTCTTGGTTGAACAATATT 724
Db 819 GAGAAATTAATGCTGTTGCTGCTGCGATACCTTCATGTTCTTGGTTGAACAATATT 878
Qy 725 TTAATCTCTGTAATCTCCCAAGATGCTGTTAGCAAAAGTCGATGCGAGAGCAAA 784
Db 879 TTAATCTCTGTAATCTCCCAAGATGCTGTTAGCAAAAGTCGATGCGAGAGCAAA 938
Qy 785 GGTGTGCAAAAGTTACGGGGGAAAGACATGCTCAGGTGAATGTTCTCTTCTCGA 844
Db 939 GGTGTGCAAAAGTTACGGGGGAAAGACATGCTCAGGTGAATGTTCTCTTCTCGA 998
Qy 845 AGGTTGAGAGTTGAGAGAGACATTCATGGAAGTACATCATTTGACCTGCGACGA 904
Db 999 AGGTTGAGAGTTGAGAGAGACATTCATGGAAGTACATCATTTGACCTGCGACGA 1058
Qy 905 GCGAGCCGATGATCTTGTATCTGACGGTGAATAGGAACAATGCACTTTATGGGCTGA 964
Db 1059 GCGAGCCGATGATCTTGTATCTGACGGTGAATAGGAACAATGCACTTTATGGGCTGA 1118
Qy 965 AGAAGGCCATGATGATGATGCTGCACTTCTAGAGGACCATCATGCTTGAAGTGTCT 1024
Db 1119 AGAAGGCCATGATGATGATGCTGCACTTCTAGAGGACCATCATGCTTGAAGTGTCT 1178
Qy 1025 TTCTCTTGATCTGCTGATGAGAGATGTTGAAACAGAGTGAACCCCTTATGATCCGAT 1084
Db 1179 TTCTCTTGATCTGCTGATGAGAGATGTTGAAACAGAGTGAACCCCTTATGATCCGAT 1238
Qy 1085 TGTGACATCTTTTGTGATGCTGATGAGAAATATGCTTCAAGCTGAGAGAAATGAGAG 1144
Db 1239 TGTGACATCTTTTGTGATGCTGATGAGAAATATGCTTCAAGCTGAGAGAAATGAGAG 1298
Qy 1145 CACATGTTTCCAACTTTGGAAGTATGTTCACTGTCACAGATCAGATGCCAAATGTA 1204
Db 1299 CACATGTTTCCAACTTTGGAAGTATGTTCACTGTCACAGATCAGATGCCAAATGTA 1358
Qy 1205 GCAATGGAATGAGAGAAATCTTCATAGGAGTGAACAGAGATGACATCTGATGCTGAG 1264
Db 1359 GCAATGGAATGAGAGAAATCTTCATAGGAGTGAACAGAGATGACATCTGATGCTGAG 1418
Qy 1265 AGGTGCTATGAGAGAAATCTTCATAGGAGTGAACAGAGATGACATCTGATGCTGAG 1324
Db 1419 AGGTGCTATGAGAGAAATCTTCATAGGAGTGAACAGAGATGACATCTGATGCTGAG 1478
Qy 1325 AGGAAGAGACATTTGTCACCATGTCACCGTGAAGTGTGAGCATGAGAGAGCAAC 1384
Db 1479 AGGAAGAGACATTTGTCACCATGTCACCGTGAAGTGTGAGCATGAGAGAGCAAC 1538
Qy 1385 CTTTATGAGGAGGATGAGATGTTGAGCAGACATGATGCTGAGGAGATGAGCT 1444
Db 1539 CTTTATGAGGAGGATGAGATGTTGAGCAGACATGATGCTGAGGAGATGAGCT 1598
Qy 1445 TGTGGAAGATGTCAGAGAGAGAGATGATGATGATGAGAGAGAGTGTTCAAAG 1504
Db 1599 TGTGGAAGATGTCAGAGAGAGAGATGATGATGATGAGAGAGAGTGTTCAAAG 1658
Qy 1505 AGTCTACTTGCACCAAGAGAGATGTCGAGCTCAAGAGAGGCTCAATGTTTCACTTCC 1564
Db 1659 AGTCTACTTGCACCAAGAGAGATGTCGAGCTCAAGAGAGGCTCAATGTTTCACTTCC 1718
Qy 1565 CGGTGTGAGCAGATGTTTGTAGAGGATGATGATGATGATGATGATGATGATGATGAT 1624
Db 1719 CGGTGTGAGCAGATGTTTGTAGAGGATGATGATGATGATGATGATGATGATGATGAT 1778
Qy 1625 GTGAGACATTTTCTCAAAGGCTTGTGTCGAACCAACATGTCAGATGCTGCAATGTTCA 1684
Db 1779 GTGAGACATTTTCTCAAAGGCTTGTGTCGAACCAACATGTCAGATGCTGCAATGTTCA 1838
Qy 1685 CCCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1744
Db 1839 CCCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1898

Qy 1745 GGTGTCCTGTTAGTGTGTTGATTCAGATCTTCAACATGTTGCTGCAATTAACGG 1804
Db 1899 GGTGTCCTGTTAGTGTGTTGATTCAGATCTTCAACATGTTGCTGCAATTAACGG 1958
Qy 1805 TGTTCCTGATTAACCCCAAAATCTTGAGAGAGTGTGAGGATTAATCTTCCAA 1864
Db 1959 TGTTCCTGATTAACCCCAAAATCTTGAGAGAGTGTGAGGATTAATCTTCCAA 2018
Qy 1865 ATTTGCTCAGCTGCGATGAGATCCATCTTGAATCAGTTCTCTCACTACTTAAT 1924
Db 2019 ATTTGCTCAGCTGCGATGAGATCCATCTTGAATCAGTTCTCTCACTACTTAAT 2078
Qy 1925 GCTTCTTGATGATGCTTTCGATGCTGCTTAATGATCTTTCGGAAGAAGTGTTCCT 1984
Db 2079 GCTTCTTGATGATGCTTTCGATGCTGCTTAATGATCTTTCGGAAGAAGTGTTCCT 2138
Qy 1985 GCTAGGCAAAATTCATCTTGAATGATCTGATTAATGCTGCTGCTGCTGCTGCTGCTGCT 2044
Db 2139 GCTAGGCAAAATTCATCTTGAATGATCTGATTAATGCTGCTGCTGCTGCTGCTGCTGCT 2198
Qy 2045 TGAATTTGGATACATGAGCCATGCTTTCCTGACCGTCAAGTGTATGCTACTTCTG 2104
Db 2199 TGAATTTGGATACATGAGCCATGCTTTCCTGACCGTCAAGTGTATGCTACTTCTG 2258
Qy 2105 CTGCTTCTGATGAGATTTGCTCCATCCCAACATTTATGTCAGATCTTCCAA 2164
Db 2259 CTGCTTCTGATGAGATTTGCTCCATCCCAACATTTATGTCAGATCTTCCAA 2318
Qy 2165 CAGGTTGTGCTGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2224
Db 2319 CAGGTTGTGCTGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2378
Qy 2225 CGTCACTACAGCCTTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 2284
Db 2379 CGTCACTACAGCCTTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 2438
Qy 2285 ATATGAGTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 2344
Db 2439 ATATGAGTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 2498
Qy 2345 GGGGATGCCCCGTGAGTATTAACGAATCTTTCAGTTGCTGAGAGAGAGGCTGCTG 2404
Db 2499 GGGGATGCCCCGTGAGTATTAACGAATCTTTCAGTTGCTGAGAGAGAGGCTGCTG 2558
Qy 2405 AAAAGCTTAATTTCTTGTGATCTTGTGTCATGCTATGATGATGATGATGATGATGAT 2464
Db 2559 AAAAGCTTAATTTCTTGTGATCTTGTGTCATGCTATGATGATGATGATGATGATGAT 2618
Qy 2465 AAGGCTTTCACCAAGAGCTGAGATTAATCTTGAATGATGATGATGATGATGATGAT 2524
Db 2619 AAGGCTTTCACCAAGAGCTGAGATTAATCTTGAATGATGATGATGATGATGATGAT 2678
Qy 2525 GAACATGCTGATGCTGCTCAACAGGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2584
Db 2679 GAACATGCTGATGCTGCTCAACAGGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2738
Qy 2585 GTTAAGTGTCAATGATGATTAACATATGTTTACCTATTTGTATCTGATTAATGATCA 2644
Db 2739 GTTAAGTGTCAATGATGATTAACATATGTTTACCTATTTGTATCTGATTAATGATCA 2798
Qy 2645 GCTATTCACCGTGTGTTGCTGAGAAATCTTGAACAAGATGATTAATGATGATGAT 2704
Db 2799 GCTATTCACCGTGTGTTGCTGAGAAATCTTGAACAAGATGATTAATGATGATGATGAT 2858
Qy 2705 GATGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2754
Db 2859 GATGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2908

RESULT 3
US-10-051-902-1
; Sequence 1, Application US/10051902

Publication No. US20020178468A1
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Hiltz, Bill
APPLICANT: Kinney, Tony
APPLICANT: Tinney, Scott
TITLE OF INVENTION: Plant Sugar Transport Proteins
FILE REFERENCE: BB-1163
CURRENT APPLICATION NUMBER: US/10/051,902
PRIORITY FILING DATE: 2002-01-17
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Microsoft Office 97
SEQ ID NO 1
LENGTH: 2824
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: unsure
LOCATION: (29)
NAME/KEY: unsure
LOCATION: (622)
NAME/KEY: unsure
LOCATION: (636)
NAME/KEY: unsure
LOCATION: (638)
NAME/KEY: unsure
LOCATION: (669)
NAME/KEY: unsure
LOCATION: (771)
NAME/KEY: unsure
LOCATION: (822)
NAME/KEY: unsure
LOCATION: (856)
NAME/KEY: unsure
LOCATION: (889)
NAME/KEY: unsure
LOCATION: (896)
NAME/KEY: unsure
LOCATION: (944)
US-10-051-902-1

Query Match 79.2% Score 2200.2; DB 14; Length 2824;
Best Local Similarity 90.2%; Pred No. 0;
Matches 2372; Conservative 0; Mismatches 253; Indels 6; Gaps 2;

133 CTTCAAGAGGCGGTGACCTCGACGATATCTTGGAGAGCAAGATGTGGGGGCTGTTTC 192
194 CTTGGCGCGCGCTGGCGGAGTGGGGGCTGATATTCGGCGGCGCATGGGGGGCGGTGA 253
193 TTGTGGCCATAGTCGCTTCATCGCATCTATTTGCAAGGGGTTGGGCAATGCGACCATCG 252
254 TGGTGGCCATCGCGGCTCTATTCGCACTTGTGCAAGGGCTGGGCAATGCGACCATTTG 313
253 CAGCTGCTGTTCTGTATATTAAGAGAAATTTCAATTGCAAAATAGGCCCATCTGGAGG 312
314 CTGAGACCTGCTCTGTAACATTAAGAGAAATTTCACTGAGAGGAGCTTGTGATGAAG 373
313 GACTAATTTGTCATGTCATCTTATTCGCGCCACCATCTGTTACTATCTCCGGCCCAT 372
374 GCTCATGCTGCGCATGTTCTTCATTTGGGGCAAGCATCATCAACATCTCGGGGGCCAA 433
373 TATCAGACTTCGATTTGGCCGACGCTATGCTTATTTCTCTTCATTTGTAATCTTTCTTA 432
434 GGGCTGACTGCGTTGGTGAAGAGGCCCATCTGCTGCTGCTGCTCTCTACTTCGTCA 493
433 GGGGCTCATCATGCTATGCTGCTCTATGCTATGCTGCTGCTGCTGCTGCTGCTGCTG 492
494 GTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 553
493 ATGATTTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 552

554 ATGGGTGGATATGCTTTGGCGGTACACTTGTCTCTCTACATCTCCGAAATCTGAC 613
553 C---TTGGAGATAGAGTTTGTGAATACACTACCAATTCAGTGG---ATCAGAG 606
614 CGCACAGAAATTTCTGGGGCTGNTGAAACAGTTGCCGACAGTTCAATGGGCTCAAGGAG 673
607 GAATGTTCTTGCATACATGATGCTGTTGGGATGTCCTGTCGATCACTCCGATCACTGGA 666
674 GGATGTTCTCTCTCACTGATGAGGTTGGGATGTCCTCATGTCGCAACCTGATTTGGA 733
667 GAATATGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 726
734 GCTCATCTGCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 793
727 ATCTCTGATATCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 786
794 ACTTCTGATATCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 853
787 TGTTCGAAAGTTACGGGGGAAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 846
854 TGNTCGAAAGCTGCGGGGAAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913
847 GGTGAGGTTGGAGAGACACTTTCATTTGAAGATCATCATTTGACCTGCGACCGAG 906
914 GTTTGGGGGCTGGTAAAGATACAGTATTTAAGTACATCATTTGACCTGCGACCGAG 973
907 CAGCGATGATCTGTTCTGACGCTGATTAAGGAAATCACTTATAGGCTGAG 966
974 CAGCGATGATCTGTTCTGACGCTGATTAAGGAAATCACTTATAGGCTGAG 1033
967 AAGGCACTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1026
1034 AAGGCACTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1093
1027 CTCTGATCTGCTGATGAGAGCATGTAACCAAGTGTACCTTATAGATTCGATTTG 1086
1094 CTCTGATCTGCTGATGAGAGCATGTAACCAAGTGTACCTTATAGATTCGATTTG 1153
1087 TGACACTTTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1146
1154 TGACACTTTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1213
1147 CATTTGTTCCAAACTTTGGAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1206
1214 CATTTGTTCCAAACTTTGGAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1273
1207 AGTGGATGAAGAGATCTTCAATAGGATGACGAGAGTACGATCTGATGCTGAGAG 1266
1274 AGTGGATGAAGAGATCTTCAATAGGATGACGAGAGTACGATCTGATGCTGAGAG 1333
1267 GTGACTATGAGAGATCTTCAATAGGATGACGAGAGTACGATCTGATGCTGAGAG 1326
1334 GTGACTATGAGAGATCTTCAATAGGATGACGAGAGTACGATCTGATGCTGAGAG 1393
1327 GGAAGCAATTTGTCACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1386
1394 GGAAGCAATTTGTCACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1453
1387 TCTTGAAGGAGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1446
1454 TCTTGAAGGAGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1513
1447 CTTGGAATTTGTCACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1506
1514 CTTGGAATTTGTCACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1573
1507 TCTTCTGCAACCAAGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1566
1574 TCTTCTGCAACCAAGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1633
1567 GTGAGGAGATGTTTGAAGGATGATGATTTGATCATGCTGCTGCTTATAGTCACT 1626
1634 GTGAGGAGATGTTTGAAGGATGATGATTTGATCATGCTGCTGCTTATAGTCACT 1693

QY 1627 CAGCACTTTCTCAAAAGGCTTGTGTAACGACGATGTCAGATGCTGCCATGCTTCAACC 1686
 DB 1694 CAGCACTTTCTCAAAAGGCTTGTGTAACGACGATGTCAGATGCTGCCATGCTTCAACC 1753
 QY 1687 CATTGAGGATGCTGCAAAAGGCTTGTGTAACGATGTCAGATGCTGCCATGCTTCAACC 1746
 DB 1754 CATTGAGGATGCTGCAAAAGGCTTGTGTAACGATGTCAGATGCTGCCATGCTTCAACC 1813
 QY 1747 GTGGCCCTGTTAGTGGGTGTGTAACGATGTCAGATGCTGCCATGCTTCAACC 1806
 DB 1814 GTGGCCCTGTTAGTGGGTGTGTAACGATGTCAGATGCTGCCATGCTTCAACC 1873
 QY 1807 TTTCTGACTATACCCCAAAATTTCTGAGAACGCTGTGGCAGTATTTCTTCCAAT 1866
 DB 1874 TTTCTGACTATACCCCAAAATTTCTGAGAACGCTGTGGCAGTATTTCTTCCAAT 1933
 QY 1867 TTGGTCTGAGCTGGGATGAGATTCATCTTGTATGATGCTTCTCACTACTACTATATG 1926
 DB 1934 TTGGTCTGAGCTGGGATGAGATTCATCTTGTATGATGCTTCTCACTACTACTATATG 1993
 QY 1927 TTTCTGACTGCTTGTGCAATGCTTGTATGATGCTTCTCACTACTACTATATG 1986
 DB 1994 TTTCTGACTGCTTGTGCAATGCTTGTATGATGCTTCTCACTACTACTATATG 2053
 QY 1987 TAGGCACATTTCCATCTTGTATGATGCTTCTCACTACTACTATATG 2046
 DB 2054 TAGGCACATTTCCATCTTGTATGATGCTTCTCACTACTACTATATG 2113
 QY 2047 ATTGGGATGCTAGCCGATGCTTGTGCAATGCTTGTATGATGCTTCTCACTACTACTATATG 2106
 DB 2114 ATTGGGATGCTAGCCGATGCTTGTGCAATGCTTGTATGATGCTTCTCACTACTACTATATG 2173
 QY 2107 GCTTGTGATGGAATTTGTGCTCCATCCCAAAATTTATGTGAGATCTTCTCAACCA 2166
 DB 2174 GCTTGTGATGGAATTTGTGCTCCATCCCAAAATTTATGTGAGATCTTCTCAACCA 2233
 QY 2167 GGGTTCGTGGCTCTGTATGCTTGTGCTTCAATCTTGTGAGATGCTTCTCAACCA 2226
 DB 2234 GGGTTCGTGGCTCTGTATGCTTGTGCTTCAATCTTGTGAGATGCTTCTCAACCA 2293
 QY 2227 TCACCTAGAGCTTCTGTGATGCTTGTGCTTCAATCTTGTGAGATGCTTCTCAACCA 2286
 DB 2294 TCACCTAGAGCTTCTGTGATGCTTGTGCTTCAATCTTGTGAGATGCTTCTCAACCA 2353
 QY 2287 ATGCACTGATGCTTGTGATGCTTGTGCTTCAATCTTGTGAGATGCTTCTCAACCA 2346
 DB 2354 ATGCACTGATGCTTGTGATGCTTGTGCTTCAATCTTGTGAGATGCTTCTCAACCA 2413
 QY 2347 GGAATGCTTGTGATGCTTGTGCTTCAATCTTGTGAGATGCTTCTCAACCA 2406
 DB 2414 GGAATGCTTGTGATGCTTGTGCTTCAATCTTGTGAGATGCTTCTCAACCA 2473
 QY 2407 AAGCCATATTTCTTGTGATGCTTGTGCTTCAATCTTGTGAGATGCTTCTCAACCA 2466
 DB 2474 AAGCCATATTTCTTGTGATGCTTGTGCTTCAATCTTGTGAGATGCTTCTCAACCA 2533
 QY 2467 GGGGTTTCAACAAGAGCTGGAATTTGTGATGCTTGTGCTTCAATCTTGTGAGATGCTTCTCAACCA 2526
 DB 2534 GGGGTTTCAACAAGAGCTGGAATTTGTGATGCTTGTGCTTCAATCTTGTGAGATGCTTCTCAACCA 2593
 QY 2527 ACATCTGCTGATGCTTGTGATGCTTGTGCTTCAATCTTGTGAGATGCTTCTCAACCA 2586
 DB 2594 ACATCTGCTGATGCTTGTGATGCTTGTGCTTCAATCTTGTGAGATGCTTCTCAACCA 2653
 QY 2587 TAAGTTGATGCTTGTGATGCTTGTGCTTCAATCTTGTGAGATGCTTCTCAACCA 2646
 DB 2654 TAAGTTGATGCTTGTGATGCTTGTGCTTCAATCTTGTGAGATGCTTCTCAACCA 2713
 QY 2647 TATTTCAAGCTGCTTGTGATGCTTGTGCTTCAATCTTGTGAGATGCTTCTCAACCA 2706
 DB 2714 TATTTCAAGCTGCTTGTGATGCTTGTGCTTCAATCTTGTGAGATGCTTCTCAACCA 2773

QY 2707 TGTATATATTCATCAATCTCAAAATTAAGAAATATGCTTCTCAAAAAA 2757
 DB 2774 TGTATATATTCATCAATCTCAAAATTAAGAAATATGCTTCTCAAAAAA 2824
 RESULT 4
 US-10-051-909-1
 ; Sequence 1, Application US/10051909
 ; Publication No. US20020199217A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Helentjaris, Tim
 ; APPLICANT: Hiez, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: B01163 US CIP
 ; CURRENT APPLICATION NUMBER: US/10/051,909
 ; PRIOR FILING DATE: 2002-01-17
 ; PRIOR APPLICATION NUMBER: 60/083,044
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO: 1
 ; LENGTH: 2824
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (29)
 ; NAME/KEY: unsure
 ; LOCATION: (622)
 ; NAME/KEY: unsure
 ; LOCATION: (636)
 ; NAME/KEY: unsure
 ; LOCATION: (638)
 ; NAME/KEY: unsure
 ; LOCATION: (669)
 ; NAME/KEY: unsure
 ; LOCATION: (771)
 ; NAME/KEY: unsure
 ; LOCATION: (822)
 ; NAME/KEY: unsure
 ; LOCATION: (856)
 ; NAME/KEY: unsure
 ; LOCATION: (889)
 ; NAME/KEY: unsure
 ; LOCATION: (896)
 ; NAME/KEY: unsure
 ; LOCATION: (944)
 ; US-10-051-909-1
 Query Match 79.2%; Score 2200.2; DB 14; Length 2824;
 Best Local Similarity 90.2%; Pred. No. 0;
 Matches 2372; Conservative 0; Mismatches 255; Indels 6; Gaps 2;
 QY 133 CTTCACAGAGCCGGTACCTCGACGATATCTTGAGACAAAGATGCGGGGCTGTTTC 192
 DB 194 CTTCACAGAGCCGGTACCTCGACGATATCTTGAGACAAAGATGCGGGGCTGTTTC 253
 QY 193 TTGTGCGCATGATGCGCTTCATCGGCAATCTATGAGAGGGGTGGGACAAATGCCAATG 252
 DB 254 TTGTGCGCATGATGCGCTTCATCGGCAATCTATGAGAGGGGTGGGACAAATGCGAATG 313
 QY 253 CAGTGTCTGTTCTGATATTAAGAAAGATTTCAATTGCAAAATGAGCCCACTGTGAGG 312
 DB 314 CAGTGTCTGTTCTGATATTAAGAAAGATTTCAATTGCAAAATGAGCCCACTGTGAGG 373
 QY 313 GACTAATTTGTGATGCTTATGCGGCGCACCATGTTACTATCAATTTCTCGGGCCAT 372
 DB 374 GACTAATTTGTGATGCTTATGCGGCGCACCATGTTACTATCAATTTCTCGGGCCAT 433
 QY 373 TATGAGACTGATGCGGCGCACCATGTTACTATTTCTCTTCAATTTCTGATTTCTCA 432

QY 2587 TAAGTTCATATGCTATTACCATATGTTTACCTATTGTTACTGATATTAAGTCAAC 2646
Db 2654 TAAGTTCATATGCTATTACCATATGTTTACCTATTGTTACTGATATTAAGTCAAC 2713
QY 2647 TATTCACGCTGTTGTTGCTGTAAGAACTTTTAGAACAAAGATGATPAATGATCTGATCTGA 2706
Db 2714 TATTCACGCTGTTGTTGCTGTAAGAACTTTTAGAACAAAGATGATPAATGATCTGATCTGA 2773
QY 2707 TGTATATATATTCATATCTCAATATTAAGAAATATGTTTCTCAAAAAA 2757
Db 2774 TGTATATATATTCATATCTCAATATTAAGAAATATGTTTCTCAAAAAA 2824

RESULT 5
US-10-051-902-7
; Sequence 7, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Alien, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Kinney, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO: 7
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: Glycine max
US-10-051-902-7

Query Match 33.9%; Score 940.2; Db 14; Length 2601;
Best Local Similarity 66.0%; Pred. No. 2e-250;
Matches 1476; Conservative 0; Mismatches 713; Indels 48; Gaps 6;

QY 177 ATGTGGGGGCTGTCTTGTGCGCATATGCGCTCCATCGGCATCTATTGAGGGGTG 236
Db 175 ATGAAGGTCGCTCTTGTGCTATTTGCGCTTCATTTGTAATTTCTCAAGATG 234
QY 237 GAAGATGCCCATGCGACCTGCTCTGTAATATAAGAGAAATTTCAATTGCAAAAT 296
Db 235 GATATGCTACCATGCGCGGGCTAATGTTACATTAAGAAACCTTCTTGGGA-- 292
QY 297 GAGCCCATGTGAGAGGACTAATTTGTCAATGTCTATGCGGCCCATCTGTTACT 356
Db 293 ---CAACTATGAAAGGCTGTGTGTGAGCATGTCTCTATTTGAGCAACGGTAATCAC 348
QY 357 ACATTCCTCGGGCATTTATCAGACTGATGAGCGGCGCCCTATGCTTATCTCTTCA 416
Db 349 ACHTCTCTGTCTATACGATTTGCTGCTGCGCGACCCATGATGATATTCATCT 408
QY 417 ATTCTACTTCTTACGCGCTCATCATGCTATGCTCTCAATGTCTATGCTCTGCTG 476
Db 409 GTGCTATATTTCTTGGGTGTTGTGATGCTGTGTCTCCCAAAATGTATGTTGTG 468
QY 477 TTGGACGCTTGTGATAGATTTGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTG 536
Db 469 TTGGAGGCTACTGATGATTTGGATTTGGCTTGTGCTGCTGCTGCTGCTGCTGCT 528
QY 537 ATTGAGAAATAGCCCTTGGAGATTAAGGTTTGTGAATACCTACACAAATTAAGT 596
Db 529 ATATCTGAACGCGCTCTGAAATTAAGGGGCTCTTAATACCTTCTCAATTAAGT 588
QY 597 GATCAGAGAGATGTTCTTGTCTATCTGATGCTGTTTGGATGTCCTGTCGCTATCA 656
Db 589 GGTCTGAGAGATGTTTGTCTGATCTATGCTTGTGGCATGTCAATGATGCTCCGCG 648

QY 657 CCCATTTGAGAAATTAAGCTTGTGTGCTGCGGATPACCTTCAATGTTCTTCTTGTGTTG 716
Db 649 CTAAGCTGAGCTCATGCTTGTGGGTTCTGTCTATTTCTTCTCTTGTATTTGATG 708
QY 717 ACAATATATTTATCTTCTGTAATCTCAAGATGCTGCTTACCAAGGTCGATGAGAG 776
Db 709 ACHATTTTTTCTTCCCGGATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 768
QY 777 GCAAAAAGGTTGCAAAAGTTACGAGGAAAGAGATGCTCAAGTAAATGCTTCT 836
Db 769 GCTAAGAGGCTGCTCAAAAGATTCGCGAAGAGAGATGCTCAAGTAAATGCTTCT 828
QY 837 CTTCTGAAAGGTTGAGGTTGAGAGAGACCTTCCATTAAGATTAATGATGAGCT 896
Db 829 CTGTTGAAGGTTCTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTGAGG 888
QY 897 GCCACCGAGGAGCGATGATCTTGTATACGCTGCTGCTGCTGCTGCTGCTGCTGCT 956
Db 889 GCTGACGATGCTGCTGATGCTGATGATGATGATGATGATGATGATGATGATGAT 948
QY 957 GGGCTGAAAGGCGCATGATGATGATGATGATGATGATGATGATGATGATGATG 1016
Db 949 GATCTCCAGAGGCTTCTTGTGTTTCAAACTGCTGCTGCTGCTGCTGCTGCTGCT 1007
QY 1017 AGTGTCTTCTCTGATCTCTGATGATGATGATGATGATGATGATGATGATGATG 1076
Db 1008 -----CTTGTCTCACCATGAGATGATGATGATGATGATGATGATGATGATG 1056
QY 1077 GATCAGATTTGATGATCTTTTGTGATGATGATGATGATGATGATGATGATGATG 1130
Db 1057 GATCTCTGCTGATGATCTTGTGATGATGATGATGATGATGATGATGATGATGAT 1116
QY 1131 GAAAGTATGAGAGACATGTTTCCAACTTTGAAATGATGATGATGATGATGATG 1190
Db 1117 GAAAGCATGAGAGACATGTTTCCAACTTTGAAATGATGATGATGATGATGATG 1176
QY 1191 CATGCAAAATGAGAGGATGAGAGAGATGATGATGATGATGATGATGATGATGAT 1250
Db 1177 CATGCAAAATGAGAGGATGAGAGAGATGATGATGATGATGATGATGATGATG 1236
QY 1251 TCTGATGCTGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1310
Db 1237 TCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1296
QY 1311 GCAAGGTCGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1370
Db 1297 ACAACAGCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1356
QY 1371 ATGAGAGGCA--AACCTTGAAGGAGGATGATGATGATGATGATGATGATGAT 1427
Db 1357 ATGAGGCTGACAGTATCTATGCAAGGATGATGATGATGATGATGATGATGAT 1416
QY 1428 GATGAGGATGAGAGCTTGTGAAATGATGATGATGATGATGATGATGATGATG 1487
Db 1417 GATGAGGATGAGAGCTTGTGAAATGATGATGATGATGATGATGATGATGATG 1473
QY 1488 GAAAGGTTTCAAAAGATCTTGTGCAACAAGAGAGATGATGATGATGATGAT 1547
Db 1474 CAAGAGGATTTAAAGATTTATTAATGATGAGAGGATGATGATGATGATGAT 1533
QY 1548 TCAATGTTTCACTTCCGATGATGATGATGATGATGATGATGATGATGATGAT 1607
Db 1534 TCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1575
QY 1608 GCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1667
Db 1576 GCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1635
QY 1668 GATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1727
Db 1636 GGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1695
QY 1728 TTTGAACCTGAGAGGAGGCTGCTGTTATGCTGATGATGATGATGATGATGAT 1787

Db 1177 CATGTAATTAATGAAACATGGGATGAGAAAGCTTAAAGAGGAAAGTGAAGACTACATG 1236
Qy 1251 TCTGATGTCGAGAGGATGACTATGAGCAATCTCAGAGCCATGCTGTCAGGAG 1310
Db 1237 TCAGATGCAACCCGCGGAGACTCCGATGATTAATTTGACAGTCTTTATCTCAGGCCAA 1296
Qy 1311 GCAACAGGTGGGAGGAGGAGGAGCAATGTCACCAATGTCACCCGCGGAGAGTCTTGGC 1370
Db 1297 ACAACAGCTTTGAAAAGAAAGCTTACCTCTCTCTCCATGAGCAATGCTTGGCAGC 1356
Qy 1371 ATGAGAGGCA--AACCTCTTAGGGAGGAGGAGGAGATGATGAGAGCACTGATATC 1427
Db 1357 ATGAGCGCTCAGAGTGTCTATGCAAGGCTGAGTGAAGAGTGTGATGATGAT 1416
Qy 1428 GGTGGGAGTGGCAGCTTGTGTAATGTCAGAGAGAGAGTGTGATGATGATGATGAT 1487
Db 1417 GGTGGTGGCTGCACTGGCAGTGAATGATGATGATGATGATGATGATGATGATGAT 1473
Qy 1488 GAGGTGGTTCGAAAGAGTCTACTGACCAAGAGGAGGAGTCTGGCTGAGAGAGGAGG 1547
Db 1474 CAGAGAGGATTAAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1533
Qy 1548 TCAATTTGTTCACTTCCGCGTGGGAGGAGTGTGATGATGATGATGATGATGATGAT 1607
Db 1534 TCCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1575
Qy 1608 GCTGCTTTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1667
Db 1576 GCTGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1635
Qy 1668 GATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1727
Db 1636 GAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1695
Qy 1728 TTTGAACTGAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1787
Db 1696 CTTGAAACAGAGGAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1755
Qy 1788 TTTGCTGATTAACAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1847
Db 1756 TTTTCAAGGATTAACAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1815
Qy 1848 GAGGATTTCTTCCAAATTTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1907
Db 1816 GAGGATTTCTTCCAAATTTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1875
Qy 1908 CTCACTACTTACTATGCTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1967
Db 1876 TTTCAACCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1935
Qy 1968 GGAAGAGGTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2027
Db 1936 GGCAGAGAGGAGTGTCTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 1995
Qy 2028 GTTGTGCTCAATCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2087
Db 1996 GTCAATGAGAGCTGTAATTTGCAATGATGATGATGATGATGATGATGATGATGATGAT 2055
Qy 2088 GTTATGCTACTTCTGCTGCTTCTTATGAGGATTTGATGATGATGATGATGATGATGATGAT 2147
Db 2056 GTTGTGCTTATTTCTGCTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 2115
Qy 2148 GCAAGATCTTTCAGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 2207
Db 2116 TCAGAGATTTTCCCAATGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2175
Qy 2208 TGAATCGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2267
Db 2176 TGAATGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2235
Qy 2268 GCGGATGTTTTCAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2327
Db 2236 GGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2295

Qy 2328 AAGTCCCTGAGCAAGAGGAGTCCCTTGAAGTATTAACCAATCTTTCAGGATGCT 2387
Db 2296 AAGGTTCCAGAAACAAAGGAGTCCCTTGAAGTATCTTGAATCTTTCGTTGGA 2355
Qy 2388 GCGAAGCAAGGAGGCTGC 2404
Db 2356 GCAAGCAAGGCTGCTTC 2372

RESULT 7
US-09-938-842A-1315
; Sequence 1315, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Mang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938, 842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227, 866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264, 647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300, 111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1315
; LENGTH: 2190
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1315

Query Match 31.5%; Score 875.8; DB 10; Length 2190;
Best Local Similarity 64.9%; Pred. No. 1.5e-232;
Matches 1449; Conservative 0; Mismatches 722; Indels 60; Gaps 8;

Qy 177 ATGCGGAGGCTGTTGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 236
Db 1 ATGAGTGAAGTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 60
Qy 237 GACAATGCAACATGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 296
Db 61 GATAACGCACTATTTGAGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Qy 297 GAGCCACTGTGAGGAGCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 356
Db 121 AATCCATCAGGAGGAGCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Qy 357 ACATTCCTCCGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 416
Db 181 ACATGCTGTGAGGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Qy 417 ATTCGATCTTCTCAGCGGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 476
Db 241 ATTCGATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Qy 477 TTGCAAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 536
Db 301 TTGCAAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Qy 537 ATTTCAGAAATAGCCCTTGGAGATTAAGGTTTGGCAATACATACATACATACATACATACAT 596
Db 361 ATATCTGAGTGAACCACTGAGATTAAGGAGTGTGATGATGATGATGATGATGATGATGATGAT 420
Qy 597 GATCAGAGAGATGTTCTGCTCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 656
Db 421 GGTCTGAGAGGAGTCTTATCTTATCTGATGATGATGATGATGATGATGATGATGATGATGAT 480

QY 657 CCCGATGAGAAATATCTGCTGCTGCGGATACCTTCAATGCTCTTCTGCTTGG 716
 DB 481 CTTACTGAGATGATGCTGCTGCTCTTCTTCACTCCCTGCTGCTTCTTCTTCTC 540
 QY 717 ACAATATTTTATCTTCTGAAATCTCCAAAGATGCTGCTTGAAGAAAGTGGCAAG 776
 DB 541 ACGGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 QY 777 GCAAAAAGTGTTCGCAAAAGTTAGCGGGGAAAGAGATGCTGAGTGAATGCTCC 836
 DB 601 GCAAAAGCGGCTTCTCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 QY 837 CTTCTGAAAGGATGAGATGAGAGAGACATTCCTCATTTGAAGATGATGATGCT 896
 DB 661 TTGGTTGAGGGCTTGGAAATGAGAGTGAACCAACATAGAGAGATATATATGCTCC 720
 QY 887 GCCACCGAGCGGATGATCTTGTACTGACGCTGATAGAGAAACAAATCACCTTAT 956
 DB 721 GCGATGAAATTAATCATGATCATATATAGCTGCTGATAGAGATCAATTAATTA 780
 QY 957 GGGGCTGAAAGAGCGCATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1016
 DB 781 GGTGCAAGAAAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 828
 QY 1017 AGTGTCTTCTCTTGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1073
 DB 829 AGCACTATGATGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 888
 QY 1074 ATGATCCGATTTGACACTTTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTG 1133
 DB 889 ATGATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 945
 QY 1134 AGTATGAGAGACATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1193
 DB 946 AGCATGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1005
 QY 1194 GCCAAATATGAGAGATGAGAGAGATCTTCTGCTGCTGCTGCTGCTGCTGCTG 1253
 DB 1006 CCAAGACATGAGATGAGAGAGATGAGAGATCTTCTGCTGCTGCTGCTGCTGCT 1065
 QY 1254 GATGCTGAGAGAGATGAGAGATGAGAGATCTTCTGCTGCTGCTGCTGCTGCT 1313
 DB 1066 GACCAT---GGAGATGATCTGAGATGATCTTCTGCTGCTGCTGCTGCTGCT 1122
 QY 1314 ACAAGTGGGAAAGGAGAGAGATGAGAGATCTTCTGCTGCTGCTGCTGCTGCT 1373
 DB 1123 ACAAGCATGAG---GAAAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1179
 QY 1374 AGAAGGCAAAACCTCTTAGGAGAGAGAGATGAGAGATCTGATGCTGCTGCTG 1433
 DB 1180 CATGAAATCAAGTCAAGAGAGATCTGAGAGAGAGAGAGAGAGAGATGAGAG 1239
 QY 1434 GAGATGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1493
 DB 1240 GAGATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1299
 QY 1494 GCTTCAAAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1553
 DB 1300 GCTTCT---GCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1329
 QY 1554 GCTTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1613
 DB 1330 GCTTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1386
 QY 1614 TTAAGTCAAGTCAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1673
 DB 1387 TTGTTAGCAACAGCTCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1446
 QY 1674 GCGATGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTT 1733
 DB 1447 GCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1503
 QY 1734 CTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1793

DB 1504 CCGAGAGCAAGCGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1563
 QY 1794 GGAATAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1853
 DB 1564 GGCATCAACGAGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1623
 QY 1854 ATCTTCTCAAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1913
 DB 1624 CTACTATGAGATGAGAGATGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1683
 QY 1914 ACCTTACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1973
 DB 1684 ACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1743
 QY 1974 AGCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2033
 DB 1744 AGGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1803
 QY 2034 TCCATCTAATTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2093
 DB 1804 TCAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1863
 QY 2094 GCTTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2153
 DB 1864 CTTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1923
 QY 2154 ATCTTCTCAACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2213
 DB 1924 ATTTTCTCAACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1983
 QY 2214 GAGATATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2273
 DB 1984 TGTGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2043
 QY 2274 GTTTCTGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2333
 DB 2044 GTTGTGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAG 2103
 QY 2334 CCGTGAACAAAGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2393
 DB 2104 CCGGAACTAAAGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2163
 QY 2394 CAAGGCTGCTG 2404
 DB 2164 CAAGCTGAGC 2174

RESULT 8
 US-09-938-842A-2254
 ; Sequence 2254, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; TITLE OF INVENTION: SAME, AND METHODS OF USE
 ; FILE REFERENCE: SCRIPI300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 2254
 ; LENGTH: 2205
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana

US-09-938-842A-2254

Query Match 25.2%; Score 700; DB 10; Length 2205;
Best Local Similarity 61.2%; Pred. No. 1.4e-183;
Matches 1371; Conservative 0; Mismatches 770; Indels 99; Gaps 11;

177 ATGTGGGGGGCTTCTTGTGGCCATGTCGCTCCATGCGCAATCTATGAGGGGCG 236
1 ATGAGAGAGGACATCTCTGCTCTGCGCCCAATCGGCAATTTCTTACAGAGATGG 60
237 GACAAATGCGCCATCGACAGCTCTGCTATATATAAGAGAAATTTCAATGCAAAAT 296
61 GACAAATGCGCCATCGACAGCTATGTTATATCAACAAGACTTGAATACCA-- 118
237 GAGCCCACTGTGGAGGACTAATTTGTCAATGTCATTTATGGGCCCACTGTTACT 356
119 ---CCCTGTTCAAGGCTTGTGCTATGTCATGATCGGTGCAACGGTCAATCAG 174
357 ACATCTCCGGGCGCATTTATCAGACTGATGGCCGACCGCCCTATGCTTCTCTCA 416
175 ACTTCTCAGAGCCCATATCTGATGCTGCGCAAGCGCCCATGCTCATTTATCATCA 234
417 ATTGTACTTCTTACGGGCTCATCATGCTATGCTCTCTATGCTATGCTCTGCTG 476
235 GTTATGTAATTCGTGCGGTTGATTAATGTTGTGCTCCCAATGTCTATGTTGTC 294
477 TTGGACGCTTGTGATGATGATTTGGATTTGGCTTGTGCTGACGCTTGTGTTAC 536
295 TTTGCTAGGCTTCTTAATGGGTTTGTGCGGGCTCGGTTACACTTGTCTCTGTTAC 354
537 ATTTGAGAAATAGCCCTTGGAGATTTAGGTTTGTGATATACATCCCAATTCAGT 596
355 ATTCTGAAACCGCTCTCCGAGATCAGAGACATTAATATCTCTCCTAGTTTCTT 414
597 GGAATAGAGAGATGTTCTTGTCTATCTGATGCTGTTGGATGCTCCCTGTGCGCATCA 656
415 GGCCTGTGAGAAATGTTTTGTCTATGCTATGTTTCACTATGCTCCGTGATGCTCC 474
657 CCCGATTTGAGAAATATGCTTGTGTGCTGCTGCGGATCTTCAATTTGTTGTTTG 716
475 CCTAGCTGAGAGCATGCTCGGTGCTCTCGATCCCTTCTCTTATTTGTTTCTC 534
717 ACAATATTTTATCTTCTGAAATCTCCAGATGCTCTTGAAGAAAGTGGATGCGAG 776
535 ACGGTTTATTTGCTCCGAGTCTCTGTTGCTGTTAGTAAAGAAAGATGAGCAG 594
777 GCAAAAAAGTGTGCAAAAGTTTACGGGGGAAAGACATGCTCAGTGAATTTGCCCTT 836
595 GCTAAGCAGATTTCTCAACAGTTATGTGGCAGAGAAATGTTACCGATGAGATGCTTTA 654
837 CTTCTCAGAGGTTGGAGTTGGAGAGACATCTTCATTTGAAGAGTACATCTGGACCT 896
655 CTAGTTGAGAGATGATATGAGAGAGAAACAAATGAGAGATCTTTAGTAACTTTG 714
897 GCGACCGAGGAGCGCATGAT---CTGTGTTACTGACGAGTAAAGAAATCAATCACT 953
715 GAGATCATGAAGGTGATGATACCTTGAACCGTTTATGAGATGACAAATGCGGCTT 774
954 TATGGGCTGAGAGAGCGCAGTCAATGCTGACCTTCAAGGAGCCCATCATGCTT 1013
775 TATGGAACCAAGAAATCAATCGTACCTGCTAACCCTGCCAGAACAA----- 826
1014 GGAAGTGTCTTCTTCTGCACTCTGTCATGAGAGCATGAGAAATATGCTCAAGCTG 1073
827 -ATGTCTCACTTGGCTACGCTCTGCGACGGAACCTTACGAAACAAAGATGATCTT 885
1074 ATGATCCGATTTGAGACATTTTGTGATGTCATGAGAAATATGCTCAAGCTGAGAGA 1133
886 AAGATTCGCTGTCATCTTTTGTGAGTCTCCAGAGAGATGCCAGAGCGCGGA 945
1134 AGTATGAGAGACATTTGTTCCAACTTTGGAATATGTTGAG-----TGTCAAGAT 1187
946 AAGCTCGAGATGGGATTTTCCCTCATTTCCGAGACATGTTCAAGTACTAGCCGATCG 1005

1188 CAGCATGCCAAAAATGACAGTGGGATGAGAGA-----ATCTTCATAGGATGAC 1238
1006 CCTACCGGTAAACCGGCTCATTTGGAAAAAGACATAGAGCCATTCACAAAGACAAAT 1065
1239 GAGAGTACGATCGATGATGATGAGAGAGTGAATATGAG-----ACAAATCTCAT 1269
1066 GATGACTATGAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1125
1290 AGCCATTTGCTGTCCAGGACGCAACAGTGGCGGAAGGGAAGACATTTGACCATGCT 1349
1126 AGCCCTTTAATGTGCGCGCAGACCAAGCATGGA---CAAGATATGATCCCATCTCT 1182
1350 CACCGTGAAGTGTCTTGGAGATGAGAGGAAACCTCTTATGAGGAGGATGATGAT 1409
1183 ACAAGTGAACCACTTTAAGATGAGAGACAGTACGATGCTTATGCA---AGGCAACGGC 1239
1410 GTGAGCAGCACTGATTTATCGGTGGGAGATGGCAGCTCTGTTGGAATGGTCAAGAGAA 1469
1240 GAAAGTACATGGAAATTTGTTGTTGGTGGATATGAGAT----- 1279
1470 GGTGAGAAATGTTAGAAAAGAGTGTTCAAAAGATCTACTTGCACCAAGAGGAGTT 1529
1280 -----ATGATATCGAAAACGATGAAATACAAAGATATTTCTTAAAGAGATGGA 1329
1530 CTTGCTCAGAAAGGGGCTCAATTTGTTCACTTCCGCTGTGTCGATGTTTGAAGGT 1589
1330 GCTAAATCTGCGCGTGTGCTGATCATCTTAATCCGAGAGTCCGATGTTG-----GGA 1383
1590 AGTATGTTTATCATGCTGCTGCTTTATGATGATGATGATGATGATGATGATGATGAT 1649
1384 GGCAGCTACATTCACGCTTCTGCTCTGATGAGATGATGATGATGATGATGATGATGAT 1443
1650 GCTGAACACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1709
1444 CATGGA-----TCCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1488
1710 TCAGTTTGAAGAAATTTGTTGAACTGAGATGAGAGGCTGCTGTTAGTGTGTTGGA 1769
1489 CCACTTGTGCTGCTCTTCTTGAACCTGAGTGAAGGCTGTTGTTGTTGTTGTTGTTG 1548
1770 ATTGATATCTTCAACAGTTTGTGTAATTAACGCTTCTGATATTAACCCCAATTT 1829
1549 ATTCAATATCGACAGGATTTTCAAGTATCAATGAGATTTCTTACATCTCTCATATT 1608
1830 CTTGAGCAAGCTGTGTGAGGATTTCTTCCAAATTTGCTCAGCTCGGATCAGCA 1889
1609 CTCGAAAGGCTGCGATGATATTTCTTTCGAGCTCGGACATGATTCATCTGCGC 1668
1890 TCCATCTTGAATCACTTCTCTCAGTACTTACTTACTTCTTCTGATTTGCTTGGCATG 1949
1669 TCATTTCTCATACAGCGGTTTAAACAAATTAATCAATGCTCCAGCAATTTGCTGCTC 1728
1950 CTGCTTATGATCTTTCGGAAGAGGTTTGTGCTAGAGCAATTTCAATCTTGATA 2009
1729 AGATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1788
2010 GCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2069
1789 GTCTCATCTGCTGCTTGTGATCAGGAGCTCATCAATCAGCAAAAGTGTGAAAGCA 1848
2070 TTGCTCTCCACCGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2129
1849 GCACTCTCCACAGGTTGTGTGCTCTTACCTTCTGCTTCTGATGAGGTTACGCTCC 1908
2130 ATCCCAACATTTTATGAGAGATCTTTCACACAGGTTGCTGCTGATATGCTC 2189
1909 ATTCAAAATCTCTCTTCTTGAATTTTCCCAAGAGTCCGTGCTCTGATGATGCC 1968
2190 ATTGCTCTTATCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2249
1969 ATATGCTATGCTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2028

QY 2250 CTGAATGCTATTTGAGCTGGCGGGTCTTTACAGCATATATGACGTGTAAGCTTGAATTC 2309
DB 2029 CTCACCTGATCGACTGAGCTAGTGTGTTTTCAGCATTTAGCTGGCGGTTTGGCTATTC 2088
QY 2310 TTTGCTTGTCTTCTTCTTAAGTCCCTGAGACAAAGGGAGTCCCTTGAGTTATTC 2369
DB 2089 TGGATCTTGTCTTCTTCTTAAGTCCCTGAGACAAAGGGAGTCCCTTGAGTTATTC 2148
QY 2370 GAATCTTTGAGCTGTTGTC 2389
DB 2149 GACTACTTGTCTTGTGAGC 2168

RESULT 9
US-10-051-902-13

; Sequence 13, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tiney, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1487
; TYPE: DNA
; ORGANISM: Trifolium aestivum
US-10-051-902-13

Query Match 24.9%; Score 691.2; DB 14; Length 1487;
Best Local Similarity 80.1%; Pred. No. 3e-181;
Matches 852; Conservative 0; Mismatches 203; Indels 9; Gaps 3;

QY 1395 GAGGCTGAGATGTGTGAGACGACCTGATATCGGTGGGAGATGCGAGCTTCTTGAAA 1454
DB 12 GAGGCTGAGGAGGAGCTGACGACGACCTGATATCGGTGGGAGATGCGAGCTTCTTGAAA 71
QY 1455 TGTGTGAG 1514
DB 72 TGTGTGAG 131
QY 1515 CACCAAG 1574
DB 132 CACCAAG 191
QY 1575 GATG---TTTTGAG 1631
DB 192 GATGACAG 251
QY 1632 CTTTCTCAAG 1688
DB 252 CTTTCTCAAG 311
QY 1689 TCTGAG 1748
DB 312 TTTGAG 371
QY 1749 GCCCTGTTAGTGTGTGAG 1808
DB 372 GATGCTGT 431
QY 1809 CTGTACTATATCCCAAAATTTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1868
DB 432 CTGTACTATATCTCTCAAAATTTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 491
QY 1869 GGTCTAGCTGGGAGATCAGATCATCTTGTATCATAGTTCTCTCACTTACTTACTT 1928

DB 492 GGCCTCAGTTAGAGATCAGATCATCTTGTATCAGTTCTCTCAGACCTTACTATCTC 551
QY 1929 CTTGACATGTGCTTTGACAG 1988
DB 552 CCAAGCATGT 611
QY 1989 GGCACATTTCCATCTTGTATGAGATCTGTATGATCTGTGTGTGTGTGTGTGTGTGTGTGTAT 2048
DB 612 GGCACATTTCCATCTTGTATGAGATCTGTATGATCTGTGTGTGTGTGTGTGTGTGTGTGTAT 671
QY 2049 TTTGAGTACAG 2108
DB 672 TTTGAGTACAG 731
QY 2109 TTTGAGTACAG 2168
DB 732 TTTGAGTACAG 791
QY 2169 GTTGTGAG 2228
DB 792 GTTGTGAG 851
QY 2229 ACTTACAG 2288
DB 852 ACTTACAG 911
QY 2289 GCACTGTATGAG 2348
DB 912 GCACTGTATGAG 971
QY 2349 ATGAG 2405
DB 972 ATGAG 1031
QY 2406 AAGGCTTATTTCTTGT 2449
DB 1032 AAGGCTTATTTCTTGT 1075

RESULT 10
US-10-051-909-13

; Sequence 13, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Heltz, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tiney, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1487
; TYPE: DNA
; ORGANISM: Trifolium aestivum
US-10-051-909-13

Query Match 24.9%; Score 691.2; DB 14; Length 1487;
Best Local Similarity 80.1%; Pred. No. 3e-181;
Matches 852; Conservative 0; Mismatches 203; Indels 9; Gaps 3;

QY 1395 GAGGCTGAGATGTGTGAGACGACCTGATATCGGTGGGAGATGCGAGCTTCTTGAAA 1454
DB 12 GAGGCTGAGGAGGAG 71
QY 1455 TGTGTGAG 1514

Db 72 TGTGGAGCCGACAGGCGAGATGGCAGAGAGAGAGGCTTCAAAAAGATCTACTTGG 131
Qy 1515 CACGAGAGAGAGGAGTCTGGCTCAAGAGAGGAGCTCAATTTTCACTTCCGGTGGG 1574
Db 132 CACGAGAGAGAGGAGTGGCGACTCAAGAGAGGAGCTCGTGTTCCTTCACTTCCGGTGGG 191
Qy 1575 GATG---TTTTAGAGGAGTGAAGTGTTCATGCTGCTGCTTTAGTGAAGTGAAGTGA 1631
Db 192 GATGCGACGAGAGAGGAGTGGGAGTGTTCATGCTGCTGCTTTAGTGAAGTGAAGTGA 251
Qy 1632 CTTTCTCAAGAGAGTCTGCTGAGCAAGCAGCATGTCATGCT---GCCATGCTTCA 1688
Db 252 CTTTACTCCAGAGATCTTATGAGAGAGCTATGGGCGCGGCTCAGCATGATTTATCCA 311
Qy 1689 TCTGAGTGAAGTCTCCAGAGGTTCAAGTTCAGTGAAGAGTGTGTTGAACTTGAAGTGAAG 1748
Db 312 TTGAGAGGAGGCTCCAGAGGTTCAATCTGAGAGAGTGTGTTGAACTTGAAGTGAAGGCT 371
Qy 1749 GCCCTGTTAGTGGTGTGTTGAGATTCAGATCCTTCAACAGTTTGTGAGATTAACGCTGT 1808
Db 372 GATGTTGCTGCTGCTGTTGAGATTCAGATGCTTCAAGAGTGTGCTGAGATTAATGAGTGT 431
Qy 1809 CTGACTATACCCCAAAATTTCTGAGCAAGCTGGGAGTGTGAGTGTCTTCTTCAAAATTT 1868
Db 432 CTGACTATACCTCCAAATTTCTGAGCAAGCTGGGAGTGTGAGTGTCTTCTTCAAAATTT 491
Qy 1869 GGTCTACGCTGGATCAGATCAGATTCATCTTGAATCAGTCTCTCAGTCACTTAATGCTT 1928
Db 492 GGGCTCAGTTCAGATCAGATCAGATTCATCTTGAATCAGTCTCTCAGTCACTTAATGCTT 551
Qy 1929 CTTTCAATGCTTGGCAGTGTGCTGATGATCTTTCGAGAGAGGTTTGTGCTGA 1988
Db 552 CCAGCATTTGGTGAAGCAGTGAAGTGAATGATGATGATGATGATGATGATGATGATGATG 611
Qy 1989 GGCAGATTTCCAACTTGTATAGATCTCTGATGATCTGATGATCTGATGATCTGATGAT 2048
Db 612 GGCAGATTTCCAACTTGTATAGATCTCTGATGATCTGATGATCTGATGATCTGATGAT 671
Qy 2049 TTGGGTACAGTCCAGATGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2108
Db 672 TTGAGTACGAGTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 731
Qy 2109 TTGCTTATGAGATTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2168
Db 732 TTGCTTATGAGATTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 791
Qy 2169 GTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2228
Db 792 GTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 851
Qy 2229 ACCTACAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2288
Db 852 ACCTACAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 911
Qy 2289 GCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2348
Db 912 GCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 971
Qy 2349 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2405
Db 972 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1031
Qy 2406 AAAGCTTATTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2449
Db 1032 ATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1075

RESULT 11
US-10-051-902-9
; Sequence 9, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:

APPLICANT: Allen, Steve
APPLICANT: Hiez, Bill
APPLICANT: Kinney, Tony
APPLICANT: Kinney, Scott
TITLE OF INVENTION: Plant Sugar Transport Proteins
FILE REFERENCE: BB-1163
CURRENT APPLICATION NUMBER: US/10/051,902
PRIOR FILING DATE: 2002-01-17
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Microsoft Office 97
SEQ ID NO 9
LENGTH: 1692
TYPE: DNA
ORGANISM: Glycine max
US-10-051-902-9

Query Match 21.4%; Score 594.6; DB 14; Length 1692;
Best Local Similarity 66.1%; Pred. No. 2.66-154;
Matches 969; Conservative 0; Mismatches 459; Indels 39; Gaps 6;

Qy 929 CGGTGATTAAGAGCAAAATCACTTATGAGGCTGAGAGAGGCGCATGATGCTGCG 988
Db 17 CAGAGAGAGAGATCAATTTAGTGTATGAGACCAAGAGAGGCGCATGCTGCTGCTAG 76
Qy 989 ACCTTCAAGGAGCCATCATGCTTGAAGTGTCTTCTTGCATCTGCTCATGAGAG 1048
Db 77 ACCTTGTCTGAGCAAAATCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 124
Qy 1049 CATGCTGAACAGGCTGATCCCTTATGATCCGATTTGATGATCACTTTTGTATGCTCA 1108
Db 125 CATGCAATTCAGAGCT---CTAGTGAACCTTATGATCCCTTGTGATGATGATCA 181
Qy 1109 TGAGATATGCTGCTGAGCTGAGAGATGAGAGAGCAATTTTCCAACTTTGAGAG 1168
Db 182 TGAGAGCTCCCAAG 229
Qy 1169 TATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1228
Db 230 TATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 289
Qy 1229 TAGGAGTGAAGAGAGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1288
Db 290 CAGAGAGGCTGATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 343
Qy 1289 TAGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1348
Db 344 GAGTCAATTTATCTCAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 400
Qy 1349 TCACCTGAGAGAGCTTGTGAGCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1408
Db 401 CCATAGTACCTGCTGAG 460
Qy 1409 TGTGAG 1468
Db 461 CACTGTGAGTGAAG 520
Qy 1469 AGGTGAG 1528
Db 521 GGGCCAG 580
Qy 1529 TCTGCTCAAG 1588
Db 581 TCTGAGTGAAG 637
Qy 1589 TAGGAGTGTGATGAG 1648
Db 638 CAGTGAAGTGTGATGAG 697
Qy 1649 TGCTGAACAG 1708
Db 698 TATGCTGAACAG 757

QY 1709 TTCAGTTGAAAGATTGTTGAACCTGAGTGAAGGCGTCCCTGTAGTGGTGTGG 1768
DB GCCAAGTTGAGATGATCTTTTGAACCTGGGGTGAAGATGATGATTTGGGGGTGG 817
QY 1769 AATTAGATCTTCAACAGTTGCTGGAATAAAGCGTGTCTGTAATTAACCCCAAT 1828
DB AATGAAATTTCTTCAAGAGTTCTCTGATATAAATGGGGTCTCTATACCCCAAT 877
QY 1829 TCTTAGAAGAGCTGTGTGAGAGTATCTTCCAAATTTGGTCTACGTGGGATGAGC 1888
DB TCTTAGAAGAGGAGGTGTGTGATCTCTTCAAGCCCTAGGCTGTGTCTACTCTTC 937
QY 1889 ATCCATCTTGAATCAAGTTCTCTCACTAATCTAATATGCTTCTGATTTGGCCAT 1948
DB ATCTTTTATTAATGAGGATGAGCAACCTGTGTATGCTCTCTGTATAGCATTTGCCAT 997
QY 1949 GCTGCTTAGATCTTCCGGAAGAGGTTTGTCTGCTAGGACAAATTCATTTGAT 2008
DB GAGGCTCATGATATTTGAGGAGAGGATTTGCTGCTCAGTACAAATCCCGTCTAAT 1057
QY 2009 AGCAATCTAGTATCCGTTGTGTCCATCTAATATGATTTGGGTACATAGCCCATGC 2068
DB AGCACTCTTCTCATATTAATGCTGGAAGTCTGTGATTTGGGATCCTCAATGAC 1117
QY 2069 TTTGCTTCCACCGTCACTGTATGCTCTAATCTGCTGCTGTATGGGATTTGCC 2128
DB ATCAATCTCAACCATATGATTTATGCTATTTCTGTTCTTTGATGGATTTGGACC 1177
QY 2129 CATCCCAATTTTATGTGAGAGATCTTTTCAACAGGTTGTGTCCTGTATTTGC 2188
DB AATCTCAATATTAATCTTGTGAGAGATCTTCCCACTGAGATTCGTGTCTGTATTC 1237
QY 2189 CATTTGACCTTATCATCTGATGAGATATCAATCTCAACCTTACCTTCTGTAT 2248
DB TATTTGACCTTATCATCTTGTGATGATATCAATCTCAACCTTACCTTCTGTAT 1297
QY 2249 GCTGATCTATATGAGTGTGCGGCTGTTTTCAGCATATATGAGTGTATGCTTATTC 2308
DB GCTGATCTATATGAGGCTTGTGATGATATCAATCTCAACCTTACCTTCTGTAT 1357
QY 2309 CTTTGTGTCTCTTCTTAAAGTCTCTGAGACAAAGGAGTCCCTTGAAGTATTTAC 2368
DB ATGGGTGTGTCTTCTTGAAGTCTCAAGAAACCAAGGAGTCCCACTGGAAGTATCAT 1417
QY 2369 CGAATCTTGTGAGTGTGCGAACA 2395
DB 1418 TGAGTTCTTCTGTGAGGACAAACA 1444

RESULT 12
US-10-051-909-9
Sequence 9, Application US/10051909
Publication No. US2002019217A1
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Helentjaris, Tim
APPLICANT: Hite, Bill
APPLICANT: Kinney, Tony
APPLICANT: Linney, Scott
TITLE OF INVENTION: Plant Sugar Transport Proteins
FILE REFERENCE: B81163 US CIP
CURRENT APPLICATION NUMBER: US/10/051,909
PRIOR FILING DATE: 2002-01-17
PRIOR APPLICATION NUMBER: 60/083,044
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Microsoft Office 97
SEQ ID NO 9
LENGTH: 1692
TYPE: DNA
ORGANISM: Glycine max
US-10-051-909-9

Query Match 21.4%; Score 594.6; DB 14; Length 1692;
Best Local Similarity 66.1%; Pred. No. 2-66-154;
Matches 969; Conservative 0; Mismatches 459; Indels 39; Gaps 6;

QY 929 CGGTATAGGAAACAAATACACTTTATAGGCTGGAAGAGCCAGTATGATGCTCG 988
DB CAGAGAAAAGATCAAAATTAATGATGACCAAGAACAGGCACTTGGTGTAG 76
QY 989 ACCCTTAAGGAGACCATCAATGCTTGAAGTGTCTTCTTGCATCTGTCATGGAG 1048
DB ACCCTTGTGAGACCAATCTGTGGC-----CTTGTATCTAGAAAAGAG 124
QY 1049 CATGATGAACAGAGTGAACCTTATGATCCGATTTGATGACATTTTGGTAGTGTCA 1108
DB CATGCAATATCAAGCACT---CTAGTGAACCTCTGTATGACCTCTTGTGTATGCA 181
QY 1109 TGAATAATGCTCAAGCTGAGAGAGATGAGAGACATTTGTTCCAACTTTGAGAG 1168
DB TGAAGAGCTCCAGAAACAGGA-----AGCACTTTTCCACATTTGGAG 229
QY 1169 TATGTTAGTGTCAAGATCAAGCTGCAAAATGAGCATGAGATGAAGAACTTTCA 1228
DB TATGTTAGTGTGAGGGAATCAAGCAAGAAATGAAGATGGATGAGAAAGCTAGC 289
QY 1229 TAGGATGAGAGGAGATAGCATCTGATGTCAGAGAGTATGAGCAATCTTCCA 1288
DB CAGAGAGGATATATTAATGCTGTAGTCT-----GATATCTGTAGCAATTTGCA 343
QY 1289 TAGCCATTTGCTGTCCAGGACAGCAAGTGTCCGAAGGAGAGACATTTGCAATGG 1348
DB GAGTCAATGATCTCAAGTCAAAACAGAGTGTGGA---TAGAGCAATCTCTCATATGC 400
QY 1349 TCACCTGAGAGTCTTGTGACATGAGAGCAAACTCTTAAAGGAGGATGAGATGG 1408
DB CCATAGTAACTTTCAGACATGAGGACAGAGTCTTTTACATGAAATTCAGAGAAACC 460
QY 1409 TGTGAGACGACATATATCGTGGGAGATGAGCAGCTTGTGAAATGATGAGAAAGGA 1468
DB CACTGTGATGATCTGAGATTTGTGTGTGTGACATGAGAAATGATTTGAAAGGA 520
QY 1469 AGGTGAGATGTGAAGAGAGGTGTTTCAAAAGATCTACTTGTGACCAAGAGAGT 1528
DB GGGCCAGATGAGAAAGAGAGGTGCTTCAAGAAATATATTTACCAAGATGGTGG 580
QY 1529 TCTGAGCTCAAGAGGCTCAATTTGTTCACTTCCGAGTGTGAGAGG 1588
DB TTTGAGATCTAGAGCTGTGTGTGTTCACTCC---TGGCGGTATTTACCAACTGA 637
QY 1589 TAGTGAATTTGATATGCTGTGCTTTAGTAAATGATGACATCTTTTCAAGAGTCT 1648
DB CAGTGAATTTGATAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 697
QY 1649 TGCTGAACACGACATGTCAGATGTCGACATGATCCCATCTGAGTGTGCAAGAG 1708
DB TATGCTCAACGCGCAGTGTGACACAGTATATTCCTCTGAACAAATTTGCAAAAG 757
QY 1709 TTAAGTTGAAAGATTTGTTGAACCTGAGTGAAGGCGTGTGTAGTGTGTGG 1768
DB GCCAAGTTGAGATATCTTTTGAACCTGGGGGTGAACATGATGATTTGGGGGTGG 817
QY 1769 AATTCAGATCTTCAAGATTTGCTGGAATAAAGGTTCTGTATATACCCCAAT 1828
DB AATCAAAATTTTGAAGATCTCTGTGATAAAGGAGTCTCTATATAGCTTCAAT 877
QY 1829 TCTTAGAAGAGTGTGTGAGTATTTCTTCAAAATTTGATGCTGAGTGTGATGAGC 1888
DB TCTTAGAAGAGCAGTGTGTGTATCTTCTTCAAGCTAGGCTTGTGTCTATCTTC 937
QY 1889 ATCCATCTTGAATCAATCTCTCACTAATCTAATGCTTCTTGTGACATTTGGCCAT 1948
DB ATCTTTTATTAATGAGGATGAGCAACCTTGTATGATGCTTCTGTATGAGCATTTGCCAT 997


```

Db      121 GGTTCCTTATCAAAACATTGACCTTAAGCTCTTCCACACATCTATTCTTATTAGGCTT 180
Qy      1910 CACTACCTTACCAATAGCTTCTTCCATGATGCTTGGCCATGCTGCTTATAGATCTTCCGG 1969
Db      181 GACACCTTGTCTGATGCTTCCAGCATTTGGCATCCCATGAGATCAATGATATGTCAGG 240
Qy      1970 AAGAAGTTTTGTGCTGATGAGCACAATTCATCTTGAATAGCATCTAGTATCTGCT 2029
Db      241 AAGAAGTTTCTTCTCTCTTCAACATCCCTGCTTGAATAGAGGCTAGCTGCTTGT 300
Qy      2030 TGTGTCATCTAATTAATTTGGATGACATGAGCCATGCTTGTCTTCCACCGTCAGT 2089
Db      301 TTATGTAATGATCTTGTGATGTCGGAACCATGATGACCTGGCTCTCAACGATCAGGT 360
Qy      2090 TATGCTTACTTCTGCTGCTGCTGCTTATAGGATTTGGTCCATCCCAATTTTATGTC 2149
Db      361 CATGCTTAATTTCTGCTTCTTGTGATGAGGTTTGGGCTTATCCCAATTAATCTCTGCG 420
Qy      2150 AGAGATCTTTCACACAGGAGTTGAGGCTCTGATTTGCCATTTGCTTACATTTG 2209
Db      421 GGAATTTTCCCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Qy      2210 GATGAGATATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2269
Db      481 GATGCGACATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Qy      2270 GGGTCTTTTCAACATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2329
Db      541 TGGATCTTCTGCGCATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Qy      2330 GGTCCCTGAGCAAGAGGAGTCCCTTGAAGTTATTAACCAATCTTTGAGTTGCTG 2389
Db      601 GGTCCCTGAGCAAGAGGAGTCCCTTGAAGTTATTAACCAATCTTTGAGTTGCTG 660
Qy      2390 GAAGCAG 2397
Db      661 AAAGCAG 668

```

```

RESULT 15
US-09-923-876-6145
; Sequence 6145, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Laligudi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Itco)
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 6145
; LENGTH: 285
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700458459H1
; NAME/KEY: unsure
; LOCATION: 3, 16, 20
; OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-6145

```

```

Query Match      9.7%; Score 269.2; DB 9; Length 285;
Best Local Similarity 97.8%; Pred. No. 3.2e-64;
Matches 271; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

Qy      1433 GGATGCGAGCTTCTCTTGAATAATGTCAGAGAAAGAGTGAATGTAGAAAGAGG 1492
Db      1 GGNATGGCAGCTTCTCTGTAATATGTCAGAGAAAGAGTGAATGTAGAAAGAGG 60
Qy      1493 TGGTTTCAAAAGAGCTTACTTGGACCAAGAGGAGTTCTTGCTCAAGAGGAGGCTCAAT 1552
Db      61 TGGTTTCAAAAGAGCTTACTTGGACCAAGAGGAGTTCTTGCTCAAGAGGAGGCTCAAT 120
Qy      1553 TGTTCATCTTCCCGGTGAGGAGTGTTTTGAAGGAGTGAATGTAGAAAGAGGCTGCTGC 1612
Db      121 TGTTCATCTTCCCGGTGAGGAGTGTTTTGAAGGAGTGAATGTAGAAAGAGGCTGCTGC 180
Qy      1613 TTTAGTAAGTCAATCAGCACTTTTCTCAAAAGAGGCTTGTGAACACGATGTAGATGC 1672
Db      181 TTTAGTAAGTCAATCAGCACTTTTCTCAAAAGAGGCTTGTGAACACGATGTAGATGC 240
Qy      1673 TGCATGAGTCAACCATCTGAGAGTACTGCCAAAGGT 1709
Db      241 TGCATGAGTCAACCATCTGAGAGTACTGCCAAAGGT 277

```

Search completed: January 5, 2004, 18:53:59
Job time : 885 secs

THIS PAGE BLANK (USPTO)

	RESULT 1				
	AY105508				
LOCUS	AY105508	2180 bp	mRNA	linear	HTC 16-OCT-2002
DEFINITION	Zea mays PCO114533 mRNA sequence.				
ACCESSION	AY105508				
VERSION	AY105508.1	GI:21208586			
KEYWORDS	HTC.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
REFERENCE	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 2180)				
AUTHORS	Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafy,W., Morjante,M. and Tingey,S.V.				
TITLE	Maize Mapping Project/Dupont Consensus Sequences for Design of Overgo Probes				
JOURNAL	Unpublished (2002)				
REFERENCE	2 (bases 1 to 2180)				
AUTHORS	Coe,E.H.				
TITLE	Direct Submission				
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA				
COMMENT	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSU, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of				

FEATURES

source

maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schinabe, Iowa State, then clones may be requested from ZMD:
www.zmd.iastate.edu.
Location/Qualifiers

1. 2180

/organism="Zea mays"

/mol_type="mRNA"

/db_xref="MaizeDB:63667"

/db_xref="taxon:4577"

/clone_id="Maize Mapping Project/Dupont Cornsensus
Library"

/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed Dupont contigs; this resource was
assembled by Dupont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

BASE COUNT 533 a 447 c 555 g 645 t

ORIGIN

Query Match 76.6%; Score 2126.2; DB 11; Length 2180;

Best Local Similarity 99.6%; Pred. No. 5.2e-242;

Matches 2131; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 619 CATCTGATGCTGTTGGAGATGCTCCGTCACCAATGAGAAATTAATGCTTG 678
DB 42 CGTCTTTAGTGGTGGATGCTCCGTCACCAATGAGAAATTAATGCTTG 101
QY 679 GGTGCTCGCATCTTCAATGCTTCTTGGTTGACAAATTTATCTTCCGAT 738
DB 102 GTGTCTCGCATCTTCAATGCTTCTTGGTTGACAAATTTATCTTCCGAT 161
QY 739 CTCAAGATGCTGTTAGCAAAAGTCCGATGCGACAGGCAAAAAGTGTTCAG 798
DB 162 CTCAAGATGCTGTTAGCAAAAGTCCGATGCGACAGGCAAAAAGTGTTCAG 221
QY 799 TAOCGGGGAAGAACATGCTCAGGATGATGCTTCTTCCAGAGGTTGAGTTG 858
DB 222 TAOCGGGGAAGAACATGCTCAGGATGATGCTTCTTCCAGAGGTTGAGTTG 281
QY 859 GAGGAGACCTTCCATTGAAAGTACATGATGAGCTTCCACAGGACGCGATGATC 918
DB 282 GAGGAGACCTTCCATTGAAAGTACATGATGAGCTTCCACAGGACGCGATGATC 341
QY 919 TTGTACTAGACGCTGTAAGAAACAAATCACTTATGAGGCTGTAAGAGCCAGTCAT 978
DB 342 TTGTACTAGACGCTGTAAGAAACAAATCACTTATGAGGCTGTAAGAGCCAGTCAT 401
QY 979 GATGCTGCTGACCTTCTAAGGAGCCATGATGCTTGGAGTGTCTTCTTGAATCTC 1038
DB 402 GATGCTGCTGACCTTCTAAGGAGCCATGATGCTTGGAGTGTCTTCTTGAATCTC 461
QY 1039 GTCATGAGGAGCTGAGCAAGAGTATCCCTTATGATCCGATTTGACATTTTGG 1098
DB 462 GTCATGAGGAGCTGAGCAAGAGTATCCCTTATGATCCGATTTGACATTTTGG 521
QY 1099 GTAGTGTCCATGAAATATGCTTCAAGCTGAGAGAGTATGAGAGCACTTTTCCAA 1158
DB 522 GTAGTGTCCATGAAATATGCTTCAAGCTGAGAGAGTATGAGAGCACTTTTCCAA 581
QY 1159 ACTTGGAGATGCTGATGCTGACAGATCAGATCCCAAAATGAGAGCTGGAGTGAAG 1218
DB 582 ACTTGGAGATGCTGATGCTGACAGATCAGATCCCAAAATGAGAGCTGGAGTGAAG 641
QY 1219 AGAATCTTCATAGGAGTGAAGAGTACGATCTGATGCTGAGAGTGAATGAGG 1278
DB 642 AGAATCTTCATAGGAGTGAAGAGTACGATCTGATGCTGAGAGTGAATGAGG 701
QY 1279 ACAATCTTCATAGGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 1338
DB 702 ACAATCTTCATAGGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 761
QY 1339 TGCACCATGCTACCGTGAAGTCTTTAGAGATGAGAGCAAAACCTCTTAAAGGAGG 1398

DB 762 TGCACCATGCTACCGTGAAGTCTTTAGAGATGAGAGCAAAACCTCTTAAAGGAGG 821
QY 1399 GTGAGATGCTGAGAGCACTGATATCGGTGGGAGATGAGAGCTTCTTGAATGCT 1458
DB 822 GTGAGATGCTGAGAGCACTGATATCGGTGGGAGATGAGAGCTTCTTGAATGCT 881
QY 1459 CAGAGAGAGAGTGAAGATGCTGAGAGAGAGAGTGTCTTCAAAAGTCTTACCTGACC 1518
DB 882 CAGAGAGAGAGTGAAGATGCTGAGAGAGAGAGTGTCTTCAAAAGTCTTACCTGACC 941
QY 1519 AAGAGAGAGTCTGAGTCAAGAGAGAGTCAATTTGTTCACTTCCGAGTGGAGATG 1578
DB 942 AAGAGAGAGTCTGAGTCAAGAGAGAGTCAATTTGTTCACTTCCGAGTGGAGATG 1001
QY 1579 TTTTGAAGGATGATGATGTTGATGATGCTGCTGTTAGTGAAGTCAAGCTTTTCT 1638
DB 1002 TTTTGAAGGATGATGATGTTGATGATGCTGCTGTTAGTGAAGTCAAGCTTTTCT 1061
QY 1639 CAAAGGCTTCTGAGAACAGAGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1698
DB 1062 CAAAGGCTTCTGAGAACAGAGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1121
QY 1659 CTGCAAAAGTCTGAGTCAAGTGAAGATTTGTTGAACCTGAGTGAAGGCTGCTGCTG 1758
DB 1122 CTGCAAAAGTCTGAGTCAAGTGAAGATTTGTTGAACCTGAGTGAAGGCTGCTGCTG 1181
QY 1759 TCGGTGTAATCAATCAATCTTCAACAGTGTCTGTAATTAACGCTGCTGTAATTA 1818
DB 1182 TCGGTGTAATCAATCAATCTTCAACAGTGTCTGTAATTAACGCTGCTGTAATTA 1241
QY 1819 CCCCACAAATCTTGAAGACAGTGTGAGAGTATCTTCCAAATTTGCTGAGCT 1878
DB 1242 CCCCACAAATCTTGAAGACAGTGTGAGAGTATCTTCCAAATTTGCTGAGCT 1301
QY 1879 CGGATCAGCATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTT 1938
DB 1302 CGGATCAGCATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTT 1361
QY 1939 GCTTTCAGATCTGCTTATGAGATCTTTCGGAAGAGTTCCTGCTGAGCAATTC 1998
DB 1362 GCTTTCAGATCTGCTTATGAGATCTTTCGGAAGAGTTCCTGCTGAGCAATTC 1421
QY 1999 CAATCTGATGATCTTCTTATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2058
DB 1422 CAATCTGATGATCTTCTTATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1481
QY 2059 TAGCCATGCTTTCCTCCACCGTCAAGTGTATCTGCTGCTGCTGCTGCTGCTGCTG 2118
DB 1482 TAGCCATGCTTTCCTCCACCGTCAAGTGTATCTGCTGCTGCTGCTGCTGCTGCTG 1541
QY 2119 GATTTGCTCCATCCCAAACTTATGAGAGATCTTTCACACAGGCTGCTGCTG 2178
DB 1542 GATTTGCTCCATCCCAAACTTATGAGAGATCTTTCACACAGGCTGCTGCTGCTG 1601
QY 2179 TCTGATGCTGATTTGCTGCTTCAATCTGATGCTGATGCTGATGCTGATGCTGATG 2238
DB 1602 TCTGATGCTGATTTGCTGCTTCAATCTGATGCTGATGCTGATGCTGATGCTGATG 1661
QY 2239 TTCTGATGCTGATGCTTATGAGATGCTGAGGCTGTTTCAAGATATATGCTGAT 2298
DB 1662 TTCTGATGCTGATGCTTATGAGATGCTGAGGCTGTTTCAAGATATATGCTGAT 1721
QY 2299 GCTTGAATCTTGTGCTGCTTCTTAAAGTCTTGAAGAGAGAGTGCCTTGG 2358
DB 1722 GCTTGAATCTTGTGCTGCTTCTTAAAGTCTTGAAGAGAGAGTGCCTTGG 1781
QY 2359 AGTATATTAAGAGTCTTCTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 2418
DB 1782 AGTATATTAAGAGTCTTCTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1841
QY 2419 TTTGATACCTTGTGCTGAGATGCTTGAAGTGAAGTGAAGAGAGAGAGAGAGAGAG 2478

Db 1842 TTGGTACCTTTTGTGTGCACTATGCACTGTAAGTAAAGAACTTGAAGGGGTTTACCA 1901
Oy 2479 AGAAGCTCGGAATTAATTGATTTGTGTAATGTTAAGGGAACGAACATCTGCTCAT 2538
Db 1902 AGAAGCTCGGAATTAATTGATTTGTGTAATGTTAAGGGAACGAACATCTGCTCAT 1961
Oy 2539 GCTCTCAACGATTAAGAAAGAGTCCCTCAATGGCAATAGAGTCTGTTAAGTTGTCAT 2598
Db 1962 GCTCTCAACGATTAAGAAAGAGTCCCTCAATGGCAATAGAGTCTGTTAAGTTGTCAT 2021
Oy 2599 GTCAATTACCAATGTTTACCTATTGTGACTGATTAATAGTCAAGCTATTCAAGCTG 2658
Db 2022 GTCAATTACCAATGTTTACCTATTGTGACTGATTAATAGTCAAGCTATTCAAGCTG 2081
Oy 2659 GTTGTGTAGAAATCTTTAGAACAAAGATGATATGATGATGATGATGATGATGAT 2718
Db 2082 GTTGTGTAGAAATCTTTAGAACAAAGATGATGATGATGATGATGATGATGATGAT 2141
Oy 2719 CAATCTCAATTAAGAAATATCTTTCTCAAAAAA 2757
Db 2142 CAATCTCAATTAAGAAATATCTTTCTCAAAAAA 2180

RESULT 2
LOCUS CC423708 911 bp DNA linear GSS 19-MAY-2003
DEFINITION PUH0R49TB ZM 0.6-1.0 KB Zea mays genomic clone ZMMBT499102,
genomic survey sequence.
ACCESSION CC423708
VERSION CC423708.1 GI:30903798
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 911)
WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Resnick
A., Frazer,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.
Maize Genomics Consortium
Unpublished
Other GSs: PUH0R49TD
CONTACT: Cathy WhiteLaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteLaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
SOURCE
1. 911
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBT499102"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"
BASE COUNT 242 a 264 c 163 g 242 t
ORIGIN
Query Match 29.2%; Score 810.4; DB 29; Length 911;
Best Local Similarity 99.3%; Pred. No. 1.1e-86;
Matches 814; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 817 TCTAGGGAATTCCTTCTTCGAGAGGTGGAGGTGGAGAGCACTTCATG 876
Db 820 TCAAGGGAATTCCTTCTTCGAGAGGTGGAGGTGGAGAGCACTTCATG 761
Oy 877 AAGAGTACATCATTTGACCTGACCGAGGAGCCGATGATCTTGTACTGACGATGATA 936
|||||

Db 760 AAGAGTACATCATTTGACCTGACCGAGGAGCCGATGATCTTGTACTGACGATGATA 701
Oy 937 AGAACAATCAACATTTATGGGCTGAAGAGGCCAGTCAATGATGATGCTGACCTCTA 996
Db 700 AGAACAATCAACATTTATGGGCTGAAGAGGCCAGTCAATGATGATGCTGACCTCTA 641
Oy 997 AGGAGCCATCATGCTTGAAGAGTGTCTTCTCTTGTGATCTGTCATGAGAGCATGTGTA 1056
Db 640 AGGAGCCATCATGCTTGAAGAGTGTCTTCTCTTGTGATCTGTCATGAGAGCATGTGTA 581
Oy 1057 ACCAGAGTATACCCCTTATGATTCGATGATGATGATGATGATGATGATGATGAT 1116
Db 580 ACCAGAGTATACCCCTTATGATTCGATGATGATGATGATGATGATGATGATGATGAT 521
Oy 1117 TGCCCTCAAGCTGAGAGAGATGAGAGCAATGTTTCCAACTTTGGAAGTATGTTCA 1176
Db 520 TGCCCTCAAGCTGAGAGAGATGAGAGCAATGTTTCCAACTTTGGAAGTATGTTCA 461
Oy 1177 GTGTCAAGATCAGCATGAGCAAAATGAGAGTGGATGAGAGAAATCTTCATAGGATG 1236
Db 460 GTGTCAAGATCAGCATGAGCAAAATGAGAGTGGATGAGAGAAATCTTCATAGGATG 401
Oy 1237 ACGAGAGTACGATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1296
Db 400 ACGAGAGTACGATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 341
Oy 1297 TGCTGTCCAGGACAGCAACAGTGGGAGAGGAGAGCAATGTCACATGTCACCGTG 1356
Db 340 TGCTGTCCAGGACAGCAACAGTGGGAGAGGAGAGCAATGTCACATGTCACCGTG 281
Oy 1357 GAAATGCTTTGAGCATGAGAGAGCAACCTCTTTAGGGAGAGGTGAGATGTTGAGCA 1416
Db 280 GAAATGCTTTGAGCATGAGAGAGCAACCTCTTTAGGGAGAGGTGAGATGTTGAGCA 221
Oy 1417 GCATGATATCGTGGGAGATGAGAGCTTGTGAAATGTCAGAGAGAGAGTGA 1476
Db 220 GCATGATATCGTGGGAGATGAGAGCTTGTGAAATGTCAGAGAGAGAGTGA 161
Oy 1477 ATGTGTAAGAGAGAGTGTTCAGAGATCTGTCAGAGAGAGAGTGTTCAGAGAGAGT 1536
Db 160 ATGTGTAAGAGAGAGTGTTCAGAGATCTGTCAGAGAGAGAGTGTTCAGAGAGAGT 101
Oy 1537 CAAGAAGAGGCTCAATGTTTCACTTCCCGTGTGCGAATGTTTGAAGGTAGT 1596
Db 100 CAAGAAGAGGCTCAATGTTTCACTTCCCGTGTGCGAATGTTTGAAGGTAGT 41
Oy 1597 TTGTACATGCTGCTGCTTGTAGTATGATGATGATGATGATGATGATGATGATGAT 1636
Db 40 TTGTACATGCTGCTGCTTGTAGTATGATGATGATGATGATGATGATGATGATGATGAT 1

RESULT 3
LOCUS BZ723844 845 bp DNA linear GSS 24-FEB-2003
DEFINITION PUCF160TD ZM 0.6-1.0 KB Zea mays genomic clone ZMMBT413323,
genomic survey sequence.
ACCESSION BZ723844
VERSION BZ723844.1 GI:28519500
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 845)
WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Resnick
A., Frazer,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.
Maize Genomics Consortium
Unpublished
CONTACT: Cathy WhiteLaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843

|||||

Fax: 301-838-0208
Email: whitejaw@tigr.org

Seq primer: TP
Class: sheared ends.

FEATURES
source Location/Qualifiers

1..845
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM0613323"
/clone_1ib="ZM 0.6-1.0 KB"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"
BASE COUNT 214 a 156 c 259 g 216 t
ORIGIN

Query Match 28.8%; Score 799.8; DB 29; Length 845;
Best Local Similarity 98.8%; Pred. No. 2e-85;
Matches 837; Conservative 0; Mismatches 7; Indels 3; Gaps 3;

QY 916 ATCTGTGTTACTGACGGTGTAAAGAACAAATCACACTTATGGCCCTGAAGAGCCAGT 975
DB 1 ATCTGTGTTACTGACGGTGTAAAGAACAAATCACACTTATGGCCCTGAAGAGCCAGT 60
QY 976 CATGATGTTGCTGACCTTCTTAAGGAGCCCATCATGCTTGGAAAGTGCTTTCTTTGCAT 1035
DB 61 CATGATGTTGCTGACCTTCTTAAGGAGCCCATCATGCTTGGAAAGTGCTTTCTTTGCAT 120
QY 1036 CTCGTCATGGAGAGCATGGTGAACCAAGAGTATGCCCTTATGGATCGATTTGACACTTT 1095
DB 121 CTCGTCATGGAGAGCATGGTGAACCAAGAGTATGCCCTTATGGATCGATTTGACACTTT 180
QY 1096 TTGCTAGTGTCCATGAGAATATGCTCAAGCTGAGAGAAATGAGACACATTTGTTTC 1155
DB 181 TTGCTAGTGTCCATGAGAATATGCTCAAGCTGAGAGAAATGAGACACATTTGTTTC 240
QY 1156 CAACTTTTGAAGATATGTTCACTGTCACAGATCAGATGCCAAATAAGACAGTGGAGT 1215
DB 241 CAACTTTTGAAGATATGTTCACTGTCACAGATCAGATGCCAAATAAGACAGTGGAGT 300
QY 1216 AAGAGATCTTCATAGGAGTACGAGAGTACGACATCTGATGGTGCAGAGGTCATTTG 1275
DB 301 AAGAGATCTTCATAGGAGTACGAGAGTACGACATCTGATGGTGCAGAGGTCATTTG 360
QY 1276 AAGCAATCTCCATAGCCCATTTGCTGCCAGGACGACAAAGTGCAGAGGAGACA 1335
DB 361 AAGCAATCTCCATAGCCCATTTGCTGCCAGGACGACAAAGTGCAGAGGAGACA 420
QY 1336 TTGCTACCATGTCACCGTGAAGTCTTTTGAAGCATGAGAAAGGCAACCTCTTTAGGGG 1395
DB 421 TTGCTACCATGTCACCGTGAAGTCTTTTGAAGCATGAGAAAGGCAACCTCTTTAGGGG 480
QY 1396 AGGGTGAAGATGCTGAGACAGACACTGATGCTGGGGGATGGAGCTGCTTGGAAAT 1455
DB 481 AGGGTGAAGATGCTGAGACAGACACTGATGCTGGGGGATGGAGCTGCTTGGAAAT 540
QY 1456 GGTGAGAGAGAGAGTGAAGATGCTGAAGAAAGAGTGTCTTCAAAAGAGTCTACTTGC 1515
DB 541 GGTGAGAGAGAGAGTGAAGATGCTGAAGAAAGAGTGTCTTCAAAAGAGTCTACTTGC 600
QY 1516 ACCAAGAGAGAGTCTCTGCTCAAGAGGGGCTCAATTGTTTCACTTCCCGGTGGGCG 1575
DB 601 ACCAAGAGAGAGTCTCTGCTCAAGAGGGGCTCAATTGTTTCACTTCCCGGTGGGCG 660
QY 1576 ATGTTTGAAGGGTGTGATGTTTGAATGCTGCTGCTTGTAGTAGTACAGTCACTT 1635
DB 661 ATGTTTGAAGGGTGTGATGTTTGAATGCTGCTGCTTGTAGTAGTACAGTCACTT 720
QY 1636 TCTCAAGAGGCTTGTGCTCAAGACCATGTCAGATCTGCGATGCTCAACCATCTGAGG 1695
DB 721 TCTCAAGAGGCTTGTGCTCAAGACCATGTCAGATCTGCGATGCTCAACCATCTGAGG 778

QY 1696 TAGCTGCCAAGAGTTCACG-CTTGGAAGATTTGTTGAACCTGAGTGAAGCGTGCCTG 1754
DB 779 TAGCTGCCAAGAGTTCACGTTTGAAGATTTGTTGAACCTGAGTGAAGCGTGCCTG 838
QY 1755 TTATGCG 1761
DB 839 TTATGCG 845

RESULT 4
B2989776 795 bp DNA linear GSS 25-MAR-2003
LOCUS PUDDY07TD_ZM_0.6-1.0_KB_Zea_mays_genomic_clone_ZM0613323,
DEFINITION genomic survey sequence.
ACCESSION B2989776
VERSION B2989776.1 GI:29229913
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 795)
Whitehead, C.A., Quackenbush, J., Van Aken, S., Uterback, T., Resnick
A., Frazer, C.M., Yuan, Y., San Miguel, P., Ma, J. and Bennetzen, J.
Maize Genomics Consortium
Other GSSs: PUDDY07TB
Unpublished
CONTACT: Cathy Whitehead
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitejaw@tigr.org
Seq primer: TP
Class: sheared ends.

FEATURES
source Location/Qualifiers

1..795
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM0613323"
/clone_1ib="ZM 0.6-1.0 KB"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"
BASE COUNT 197 a 238 c 143 g 217 t
ORIGIN

Query Match 26.4%; Score 732; DB 29; Length 795;
Best Local Similarity 99.3%; Pred. No. 2.1e-77;
Matches 735; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 817 TCTAGGAGATGTCCTCTTCTTGAAGGTTGAGGTTGAGAGACACTTCATTTG 876
DB 740 TCAAGGTGAATTCCTCTTCTTGAAGGTTGAGGTTGAGAGACACTTCATTTG 681
QY 877 AAGAGTACATCATTTGACCTGCCACGAGCAGCCGATGATCTTGTACTGACGGTATA 936
DB 680 AAGAGTACATCATTTGACCTGCCACGAGCAGCCGATGATCTTGTACTGACGGTATA 621
QY 937 AGGAACAAATCACCTTTATGGGCTTGAAGAGCCGATGATGATGCTGACCTTTCTA 996
DB 620 AGGAACAAATCACCTTTATGGGCTTGAAGAGCCGATGATGATGCTGACCTTTCTA 561
QY 997 AGGAGCCATCATGCTTGAAGTGTCTTCTGTCATCGTCATGGAGAGATGGTA 1056
DB 560 AGGAGCCATCATGCTTGAAGTGTCTTCTGTCATCGTCATGGAGAGATGGTA 501
QY 1057 ACCAGAGTGAACCTTATGATCCGATTTGATGACACTTTTGTGATGTCATGAGATA 1116
DB 500 ACCAGAGTGAACCTTATGATCCGATTTGATGACACTTTTGTGATGTCATGAGATA 441

ORIGIN	Query Match	21.4%	Score 594	DB 9	Length 594	Best Local Similarity 100.0%	Pred. No. 4.5e-61	Matches 594	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1795	GAATAAACGGGTTCTGTACTATATCCCAACAATTTCTTGAGCAAGCTGGTGGCAGTTA	185									
Db	594	GAATAAACGGGTTCTGTACTATATCCCAACAATTTCTTGAGCAAGCTGGTGGCAGTTA	535									
QY	1855	TTCTTTTCCAAATTTGGTCTCAGCTCGGCATCAGATCCATCTTGATCATGTTCTCTACTA	1914									
Db	534	TTCTTTTCCAAATTTGGTCTCAGCTCGGCATCAGATCCATCTTGATCATGTTCTCTACTA	475									
QY	1315	CCTTACTAATGCTTCCTTGACATTTGGCTTGGCAAGCTGATATGAGATCTTCCGGAAGA	1974									
Db	474	CCTTACTAATGCTTCCTTGACATTTGGCTTGGCAAGCTGATATGAGATCTTCCGGAAGA	415									
QY	1975	GGTTTTCGCTGCTAGGACAAATTCGAATCTTGATAGCATCTTAGTTATCCTGGTTGGT	2034									
Db	414	GGTTTTCGCTGCTAGGACAAATTCGAATCTTGATAGCATCTTAGTTATCCTGGTTGGT	355									
QY	2035	CCAACTCTAATTTGATTTGGGTCACCTAGCCCATGCTTTGCTCTCCACCGTCAGTGTATCG	2094									
Db	354	CCAACTCTAATTTGATTTGGGTCACCTAGCCCATGCTTTGCTCTCCACCGTCAGTGTATCG	295									
QY	2095	TCTACTCTGCTGCTTGATGGATTTGGTCCCATCCCAACATTTTATGTGCAGAGA	2154									
Db	294	TCTACTCTGCTGCTTGATGGATTTGGTCCCATCCCAACATTTTATGTGCAGAGA	235									
QY	2155	TCTTTCACACGGGTTCTGTGCTCTGTATATGCAATTTGTCCTTACATTTCTGAGATCG	2214									

Db	Accession	Version	Keywords	Source	Organism	Result
2215	GAGATATCANCCTGACCTTACGACCTTCCTGTCGATGCTGAAGCTATTGGACTTGGCGGGTG					227
174	GAGATATCANCCTGACCTTACGACCTTCCTGTCGATGCTGAAGCTATTGGACTTGGCGGGTG					115
2275	TTTTCAGCATATATGACAGTGGTATGCTGATTTCTTGTGTGGTCTTCCCTTAAGGTC					2334
114	TTTTCAGCATATATGACAGTGGTATGCTGATTTCTTGTGTGGTCTTCCCTTAAGGTC					55
2335	CTGAGACAAAGGGGATGCCCTTGAGGTATTATCCGAATTTTTCAGATTGGTG					2388
54	CTGAGACAAAGGGGATGCCCTTGAGGTATTATCCGAATTTTTCAGATTGGTG					1
B2819234						
LOCUS	PUCSDY07TB_ZM_0.6_1.0_KB_Zea mays genomic clone ZMMBRT362B13,					
DEFINITION	genomic survey sequence.					
ACCESSION	B2819234					
VERSION	B2819234.1					
KEYWORDS	GSS.					
SOURCE	Zea mays					
ORGANISM	Zea mays					
REFERENCE	Eukaryote: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD clade: Panicoideae; Andropogoneae; Zea.					
AUTHORS	1 (bases 1 to 817) Whitelaw, C.A., Quackenbush, J., Van Aken, S., Uteirback, T., Resnick, A., Frazer, C.M., Yuan, Y., San Miguel, P., Ma, V. and Bennetzen, J.					
JOURNAL	Maize Genomics Consortium					
COMMENT	Unpublished Contact: Cathy Whitelaw					
TICK	9712 Medical Center Drive, Rockville, MD 20850, USA					
	Tel: 301-838-5843					
	Fax: 301-838-0208					
	Email: whitelaw@igf.org					
	Seq primer: TR					
	Class: sheared ends.					

FEATURES
source
Location/Qualifiers
1. 817
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBT362B13"
/clone_lib="ZM_0.6-1.0 kb"
/note="vector: PCR4-JOFO; Site 1: EcoRI; 0.6-1.0 kb high
cor selected genomic DNA library"
BASE COUNT 232 a 145 c 203 g 237 t
ORIGIN

Query Match 19.9%; Score 553.2; DB 29; Length 817;
Best Local Similarity 99.5%; Pred. No. 2.4e-56;
Matches 555; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

817 TCTCAGTGGAATTTGCTTCTTCTCGAAGGTTGAGGTTGAGAGACATTCCTCATTTG 876
260 TCACAGGTGATTTGCTTCTTCTCGAAGGTTGAGGTTGAGAGACATTCCTCATTTG 319
877 AAGATGATCATTTGAGGCTGACCGAGGAGCGGATGATCTTGTACTGACGGTGATA 936
320 AAGATGATCATTTGAGGCTGACCGAGGAGCGGATGATCTTGTACTGACGGTGATA 379
937 AGGACCAATCATCTTTATGAGGCTGAGAGAGCGGATGATGTTGTCACCTTTCTA 996
380 AGGACCAATCATCTTTATGAGGCTGAGAGAGCGGATGATGTTGTCACCTTTCTA 439
997 AGGACCAATCATCTTTATGAGGCTGAGAGAGCGGATGATGTTGTCACCTTTCTA 1056
440 AGGACCAATCATCTTTATGAGGCTGAGAGAGCGGATGATGTTGTCACCTTTCTA 499
1057 AGGACCAATCATCTTTATGAGGCTGAGAGAGCGGATGATGTTGTCACCTTTCTA 1116
500 AGGACCAATCATCTTTATGAGGCTGAGAGAGCGGATGATGTTGTCACCTTTCTA 559
1117 TGCCTCAAGCTGAGAGAGTATGAGAGACATTTGTTCCAACTTTGGAAGTATGTTCA 1176
560 TGCCTCAAGCTGAGAGAGTATGAGAGACATTTGTTCCAACTTTGGAAGTATGTTCA 619
1177 GGTGACAGATCAGATGCGCAAAATGAGACATGAGATGAGATGAGATGAGATGAGATG 1236
620 GGTGACAGATCAGATGCGCAAAATGAGACATGAGATGAGATGAGATGAGATGAGATG 679
1237 ACGAGAGTACGATCTGATGCTGAGAGAGTATGAGAGACATTCCTCATGAGCCAT 1296
680 ACGAGAGTACGATCTGATGCTGAGAGAGTATGAGAGACATTCCTCATGAGCCAT 739
1237 TGCCTCAAGCTGAGAGAGTATGAGAGACATTTGTTCCAACTTTGGAAGTATGTTCA 1356
740 TGCCTCAAGCTGAGAGAGTATGAGAGACATTTGTTCCAACTTTGGAAGTATGTTCA 799
1357 GAAAGTCTTGGAGCATGA 1374
800 GAAAGTCTTGGAGCATGA 817

RESULT 7
LOCUS B0752737 829 bp mRNA linear EST 23-JUL-2002
DEFINITION WHB4118_E11_122Z5 wheat salt-stressed root cDNA library Triticum
aestivum cDNA clone WHB4118_E11_122, mRNA sequence.
ACCESSION B0752737
VERSION B0752737.1 GI:21930519
KEYWORDS EST
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 829)
AUTHORS Anderson, O.D., Akhunov, E., Chao, S., Crossman, C., Deal, K., Dvorak, J.

TITLE
JOURNAL
COMMENT
Lazo, G.R., Pham, J., Rausch, C.J., Wilson, C. and Woo, J.
The structure and function of the expressed portion of the wheat
genomes - Salt-stressed root cDNA library
Unpublished
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 510595773
Fax: 5105595818
Email: oanderson@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: SK primer
FEATURES
source
Location/Qualifiers
1. 829
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHB4118_E11_122"
/issue_type="Roots"
/dev_stage="Full tillering"
/lab_host="E. coli SOLR"
/clone_lib="Wheat salt-stressed root cDNA library"
/note="vector: Lambda Uni-ZAP XR, excised phagemid
pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; Hydropic
plants grown to full tillering stage were treated with 150
mM NaCl for either 12 hours or 7 days. Root tissues of the
plants subjected to both types of treatment were collected
separately at University of California, Davis (E. Akhunov
and K. Deal in U. Dvorak's lab). Total RNA was prepared
separately from the two samples (12h and 7day treatments),
and equal amount of RNA was then pooled. PolyA RNA was
purified from the pooled RNA, a cDNA library was made, and
the cDNA clones were in vivo excised to give pBluescript
SK(-) phagemids in U. Dvorak's lab (E. Akhunov, J. Dvorak)
at the University of California, Davis. Colony plating,
plasmid DNA preparations and DNA sequencing were performed
in the OD Anderson lab (all other authors)."
BASE COUNT 177 a 199 c 215 g 238 t
ORIGIN

Query Match 19.1%; Score 531; DB 13; Length 829;
Best Local Similarity 80.6%; Pred. No. 1e-53;
Matches 647; Conservative 0; Mismatches 150; Indels 6; Gaps 2;

1466 GAAAGTGAAGATGTTGAAAGAGAGGTTTCAAAAGATCTTACACCAAGAGGG 1525
27 GAAAGTGAAGATGTTGAAAGAGAGGTTTCAAAAGATCTTACACCAAGAGGG 86
1526 AGTTCTGCTCAAGAGAGGCTCAATTTGTTCACTTCCGCGTGCGCATG--TTT 1582
87 GGTGGCGCACTCAAGAGAGGCTCTGTTGTTCACTTCCGCGTGATGACACCA 146
1583 TGAGGTGATGATTTGTTGATGCTGCTTTAGTAAAGTACAGACATTTTCTCAA 1642
147 AGGGGCGAGTGGTTTATACGCTGCTGTTGGTAAACCACTGCTCTTACTCAA 206
1643 GGGTCTGCTCAACACGAT--GTGAGATGCTGCTGATGAGGCGTCTGTTAGT 1759
207 GATCTTATGAAAGGCTATGCGGCGGCTGACGATGATCATTCATGAGAGGAGC 266
1700 TGCAGAGTTTACGTTGAAAGATTTGTTGAACCTGAGTGAAGGCGTCTGTTAGT 1759
267 TCCAGAGTTTCAATCTGAAAGATCTGTTGAACCTGAGTGAAGGCGTCTGTTAGT 326
1760 CGGTGTTGATTTGATCTTCAACAGTTTGTCTGGAATTAAGGTTTCTGTTACTATAC 1819
327 CGGGTGTGAATTCAGATGCTTCAAGAGTTTGTCTGGAATTAAGGTTTCTGTTACTATAC 386
1820 CCAAGATTTCTTGAAGAGGAGTGGAGTATTTCTTCAAAATTTGATGCTCAGCTC 1879

Db 387 TCCCAAAATTCGAGCAGAGCTGTGTGCGGTTCTTTCCAACTTGGCCTCAGTTC 446
 Qy 1880 GGCATCAGATCCATCTTGATGATGTTCTCTCACTACTAATGCTTCTTGATGG 1939
 Db 447 AGCATCAGAGCTTCACTTGATGATGTTCTCTCACTACTAATGCTTCTTGATGG 506
 Qy 1940 CTTTCCATGCTGCTTATGATCTTCCGGAAGAGTTTGTCTAGGCAATTC 1999
 Db 507 TGTACCCATGAGACTTATGATATATCGAAGAGTTTCTTCTAGGCAATTC 566
 Qy 2000 AATCTTGATGATCTTATGATCTTCCGGAAGAGTTTGTCTAGGCAATTC 2059
 Db 567 CATCTGATGATGATCTTATGATCTTCCGGAAGAGTTTGTCTAGGCAATTC 626
 Qy 2060 AGCCATGCTTGTCTCTCCAGCTGATGATGATCTTCCGGAAGAGTTTGTCTAGGCAATTC 2119
 Db 627 GGCACAGCTGCTGCTCTCCAGCTGATGATGATCTTCCGGAAGAGTTTGTCTAGGCAATTC 686
 Qy 2130 ATTTGCTCCATCCCAACATTTTATGTCAGAGATCTTCCGGAAGAGTTTGTCTAGGCAATTC 2179
 Db 687 CTTTGGCCCAATCCCAACATTTTATGTCAGAGATCTTCCGGAAGAGTTTGTCTAGGCAATTC 746
 Qy 2180 CTGTATGCTGCTTGTCTCTCACTACTAATGCTTCTTGATGG 2239
 Db 747 CTGATGCTGCTTGTCTCTCACTACTAATGCTTCTTGATGG 806
 Qy 2240 TCCCTGATGCTGATGCTATG 2262
 Db 807 GCTGTGATGCTGATGCTATG 829

RESULT 8
 B1992011/c 545 bp mRNA linear EST 24-OCT-2001
 LOCUS 1020055D09.x4 1020 - Unigene II from Maize Genome Project Zea mays
 DEFINITION cDNA, mRNA sequence.
 ACCESSION B1992011
 VERSION B1992011.1 GI:16377584
 KEYWORDS EST.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 545)
 REFERENCE 1 Walbot, V.
 AUTHORS Walbot, V.
 TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
 JOURNAL Unpublished
 COMMENT Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 1020055 row: D column: 09.
 FEATURES
 source
 1..545
 /organism="Zea mays"
 /mol_type="mRNA"
 /db_xref="taxon:4577"
 /clone_lib="1020 - Unigene II from Maize Genome Project"
 /note="This library represents the unique genes found in
 the second round of EST sequencing at Stanford University
 for the maize genome project. Sequences are present from
 libraries 947 and 949. Contigs were assembled using
 ZDBAssembler and 2 representatives from each contig were
 selected for the Unigene set. All singlets were also
 selected."

BASE COUNT 182 a 120 c 101 g 142 t
 ORIGIN

Query Match 18.9%; Score 525; DB 12; Length 545;
 Best Local Similarity 98.9%; Pred. No. 6,5e-53;
 Matches 539; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 2178 CTCTGATGCTGATGCTGCTTCTTCACTGATGCGAGATATCATCTGCTACAC 2237
 Db 545 CTCTGATGCTGATGCTGCTTCTTCACTGATGCGAGATATCATCTGCTACAC 486
 Qy 2238 CTTCTGATGCTGATGCTGCTTCTTCACTGATGCGAGATATCATCTGCTACAC 2297
 Db 485 CTTCTGATGCTGATGCTGCTTCTTCACTGATGCGAGATATCATCTGCTACAC 426
 Qy 2298 TGTCTGATGCTGATGCTGCTTCTTCACTGATGCGAGATATCATCTGCTACAC 2357
 Db 425 TGTCTGATGCTGATGCTGCTTCTTCACTGATGCGAGATATCATCTGCTACAC 366
 Qy 2358 GAGGTATGCTGATGCTGCTTCTTCACTGATGCGAGATATCATCTGCTACAC 2416
 Db 365 GAGGTATGCTGATGCTGCTTCTTCACTGATGCGAGATATCATCTGCTACAC 306
 Qy 2417 TCTTGTGATGCTGATGCTGCTTCTTCACTGATGCGAGATATCATCTGCTACAC 2476
 Db 305 TCTTGTGATGCTGATGCTGCTTCTTCACTGATGCGAGATATCATCTGCTACAC 246
 Qy 2477 CAAGAGCTCGAGATATGCTGCTTCTTCACTGATGCGAGATATCATCTGCTACAC 2536
 Db 245 CAAGAGCTCGAGATATGCTGCTTCTTCACTGATGCGAGATATCATCTGCTACAC 186
 Qy 2537 ATGCTCTCAAGAGATATGCTGCTTCTTCACTGATGCGAGATATCATCTGCTACAC 2596
 Db 185 ATGCTCTCAAGAGATATGCTGCTTCTTCACTGATGCGAGATATCATCTGCTACAC 126
 Qy 2597 ATGCTCTCAAGAGATATGCTGCTTCTTCACTGATGCGAGATATCATCTGCTACAC 2656
 Db 125 ATGCTCTCAAGAGATATGCTGCTTCTTCACTGATGCGAGATATCATCTGCTACAC 66
 Qy 2657 TGTCTGATGCTGATGCTGCTTCTTCACTGATGCGAGATATCATCTGCTACAC 2716
 Db 65 TGTCTGATGCTGATGCTGCTTCTTCACTGATGCGAGATATCATCTGCTACAC 6
 Qy 2717 TTCAA 2721
 Db 5 TTCAA 1

RESULT 9
 BM325261 572 bp mRNA linear EST 04-JAN-2002
 LOCUS PIC1.42.B1.b1 A002 Pathogen-infected compatible 1 (PIC1) Sorghum
 DEFINITION bicolor cDNA, mRNA sequence.
 ACCESSION BM325261
 VERSION BM325261.1 GI:18064398
 KEYWORDS EST.
 SOURCE Sorghum bicolor (sorghum)
 ORGANISM Sorghum bicolor
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoidae; Andropogoneae; Sorghum.
 1 (bases 1 to 572)
 REFERENCE 1 Cordonnier-Pratt, M.-M., Gingle, A., Fang, G.C., Dean, R., Wing, R.,
 Sudman, M., and Pratt, L.H.
 An EST database from Sorghum: plants infected with a compatible
 pathogen
 JOURNAL Unpublished
 COMMENT Contact: Cordonnier-Pratt MM
 Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 583 0210
 Email: mmpratt@uga.edu
 Sequences have been trimmed to exclude PolyA, vector, and regions
 below Phred quality 16. The threshold for highest quality sequence

High quality sequence stop: 664
POLYA=Yes.
Location/Qualifiers
1. 664
/organism="Sorghum bicolor"
/mol_type="mRNA"
/cultivar="BTx623"
/db_xref="taxon:4558"
/note="Organ: Developing preanthesis panicle 1 (IP1)"
pbluescript II SK(-) from lambda Zap II; Site 1: XhoI;
Site 2: EcoRI; The library was made from poly-A RNA in the
cloning vector lambda Zap II. Clones to be sequenced were
prepared by mass excision."

BASE COUNT 165 a 141 c 144 g 214 t
ORIGIN

Query Match 16.9%; Score 469.6; DB 12; Length 664;
Best Local Similarity 87.2%; Pred. No. 2e-46;
Matches 586; Conservative 0; Mismatches 74; Indels 12; Gaps 6;

2077 CCACGTCAGTGTATTCGTTCTACTTCTGCTGCTTGTATGGATTGGTCCATCCCA 2136
1 CCACAGCAGTGTCTATCTACTTCTGCTGCTTGTATGGATTGGTCCATCCCA 60
2137 ACATTATAGTCAGAGATCTTTCCACACAGGGTGGGCTCTGTATTGCAATTTGG 2196
61 ACATTCTATGTCAGAGATCTTTCCACACAGGGTGGGCTCTGTATTGCAATTTGG 120
2197 CCTTACATTCGTGATCGAGATATCATCTGACCTACAGCCTTCCTGTATGCTGAAG 2256
121 CTTGACGTTTGGATTGGAAGACATCATTTGACCTACAGCCTTCCTGTATGCTGAAG 180
2257 CTATTGACGTGGGGTGTTCACCATATATGACGTGTATGCTTGTATTTCTTTGCT 2316
181 CTATTGACGTGACAGGTGTTTGGCATATATGACGTGTATGCTTGTATTTCTTTGCT 240
2317 TCGTCTTCCTTAAGTCCCTGAGACAAAGGGGATCCCTTGAGGTTTATCCGAATCT 2376
241 TTGCTCTCTTAAGTCCCTGAGACAAAGGGGATCCCTTGAGGTTTATCCGAATCT 300
2377 TTGCAATGTCGAGCAAGCGGCTGCAAAAAGCTTAATTTCTTTGTAACCTTTGTGTC 2436
301 TTGCAATGTCGAGCAAGCGGCTGCAAAAAGCTTAATTTCTTTGTAACCTTTGTGTC 360
2437 AAC-TATTGCACTGTAACTTGAAGAGGGTTCACCAAGAGCTGGAAGTA 2495
361 AACTTTTGGCACTGGAATTAATTAATTTGAAGGATTCACCAAGAGCTGGAAGTA 417
2436 CTTTGGATTGTGTAATATTTTAAGGAAACGAATCTGCTATGCTCTCAAAAGGTAA 2555
418 CTTTGGATTGTGTAATATTTTAAGGAAACGAATCTGCTATGCTCTCAAAAGGTAA 476
2556 AAAGAGTCCCTCAATGCAAAATAGAGTGTGTAATGTCATGTCATTACCATATGT 2615
477 CAAGACTCCCTGATGCAAGAGAGAGTGTGTAATGTCATGTCATTACCATATGT 536
2616 -TTACCTATTGTAATTAATTAAGTCAAGCTATTCACAGCTGTGTTAGTAATTC 2674
537 GTTACCTATTGTAATTAATTAAGTCAAGCTATTCACAGCTGTGTTAGTAATTC 593
2675 TTTAAGCAAGAGATATATGATCTGATCTGATTTATTAATTTCAATCTCAATTAAG 2734
594 TTTAAGCAAGAGATATATGATCTGATCTGATTTATTAATTTCAATCTCAATTAAG 650
2735 AAAATATGCTTT 2746
651 AAAATATGCTTT 662

RESULT 12
LOCUS BG356047 653 bp mRNA linear EST 17-DEC-2001

2149 CAGAGATCTTTCCAAACGAGGTTTCGTGCTCTGTATTCGCAATTTGCTTACATCT 2208
121 CAGAGATCTTTCCAAACGAGGTTTCGTGCTCTGTATTCGCAATTTGCTTACATCT 180
2209 GGATGCGAGATATCATCTGACCTTACAGCCTTCTGTATGCTGAATGCTATTTGA 2268
181 GGATGCGAGATATCATCTGACCTTACAGCCTTCTGTATGCTGAATGCTATTTGA 240
2269 CGGGGTTTTCAGCATATATGAGTGTATGCTGATTTCTTGTGTTGCTTCTTCTTA 2328
241 CAGGTTGTTTTCAGCATATATGAGTGTATGCTGATTTCTTGTGTTGCTTCTTCTTA 300
2329 AGGTCCTGAGCAAAAGGAGATGCCCCCTTGAGGTTATTCACCAATTTCTTGCA 2388
301 AGGTCCTGAGCAAAAGGAGATGCCCCCTTGAGGTTATTCACCAATTTCTTGCA 360
2389 CGAGCAAGCGGCTGCAAAAAGCTTAATTTCTTTGTAACCTTTGTGTCAC -TATTGCA 2447
361 CGAGCAAGCGGCTGCAAAAAGCTTAATTTCTTTGTAACCTTTGTGTCAC -TATTGCA 420
2448 TGTATGTTAGAACTTGAAGGGGTTTCACCAAGAGCTGGAATTTACTTTGATTTGT 2507
421 TGTATGTTAGAACTTGAAGGGGTTTCACCAAGAGCTGGAAGCTTGGGA - -ACTTTGATTTGT 477
2508 GTAATGTTAGAAAGCAAGACATCTGCTCATGCTCTCTCAACGCTTAAAGAGTCCCTC 2567
478 ATTAATGCGAGAGAAAGATATCTGCTCATGCTCTCTCAACGCTTAAAGAGTCCCTC 536
2568 AATGCAATAGAGAGTCTTGAAGTGTCAATGTCATTTACATATGTT -TTAATTTG 2626
537 GATGCAAAAGCAAGTGTGTTAGTGTCAATGTCATTTACATATGTTTGTACTATTTG 596
2627 TACTGTATTAATGTCAGCTATTCACGCTGCTTGTCTGTAATTTCTTTGAACAAG 2686
597 TACTGTATTAATGTCAGCTATTCACGCTGCTTGTCTGTAATTTCTTTGAACAAG - 652
2687 ATGATATGATCTGATCTGATGTTTAAATTTCAATCTCAATATAA 2733
653 -GATATGATCTGATCTGATGTTTAAATTTCAATCTCAATATAA 697

RESULT 11
LOCUS BI074258 664 bp mRNA linear EST 20-JUN-2001
DEFINITION IP1.13.H11.g1.A002 Immature panicle 1 (IP1) Sorghum bicolor cDNA,
mRNA sequence.
ACCESSION BI074258
VERSION BI074258.1 GI:14512915
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Sorghum.
1 (bases 1 to 664)
Klein,R.R., Cordonnier-Pratt,M.-M., Gingle,A., Sudman,M. and Pratt
L.H.
An EST database from Sorghum: developing preanthesis panicles
Unpublished
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Km. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for high quality sequence is
20. Three-prime sequences, which are obtained with PolyTmIX or T7
sequencing primer, are presented as the reverse complement.
Seq primer: T7
High quality sequence start: 112

DEFINITION EML 19 F09.g1_A002 Embryo 1 (EML) Sorghum bicolor cDNA, mRNA
 sequence.
 ACCESSION BJ356047
 VERSION BJ356047.2 GI:15724666
 KEYWORDS EST.
 SOURCE Sorghum bicolor (sorghum)
 ORGANISM Sorghum bicolor
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Sorghum.
 REFERENCE 1 (bases 1 to 653)
 Reid, S.P., Cordonnier-Pratt, M.-M., Gingle, A. and Pratt, L.H.
 An EST database from Sorghum: developing embryos
 JOURNAL Unpublished
 On Mar 6, 2001 this sequence version replaced gi:13238033.
 CONTACT Cordonnier-Pratt MM
 Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 TEL: 706 542 1860
 FAX: 706 583 0210
 EMAIL: mmpratt@uga.edu
 SEQUENCES have been trimmed to exclude PolyA, vector and regions
 below Phred quality 16. The threshold for highest quality sequence
 is 20. Three-prime sequences, which are obtained with PolyMix or
 T7 sequencing primer, are presented as the reverse complement.
 SEQ PRIMER: PolyMix
 High quality sequence start: 38
 High quality sequence stop: 612
 POLYA=No.

FEATURES
 SOURCE 1..653
 Location/Qualifiers
 /organism="Sorghum bicolor"
 /mol_type="mRNA"
 /db_xref="taxon:4558"
 /clone_lib="Embryo 1 (EML)"
 /note="Organ: Embryos germinated for 24 hr; Vector:
 Bluescript II from Lambda Zap II; Site 1: XhoI; Site 2:
 EcoRI; The library was made from polyA-RNA in the cloning
 vector Lambda Zap II. Clones to be sequenced were
 prepared by mass excision."
 BASE COUNT 153 a 142 c 149 g 209 t
 ORIGIN

Query Match 16.8%; Score 466.2; DB 10; Length 653;
 Best Local Similarity 86.8%; Pred. No. 5e-46; Indels 9; Gaps 5;
 Matches 572; Conservative 0; Mismatches 78; Indels 9; Gaps 5;

QY 2032 TGTCATCTAATGATTTGGGTACACAGCCCATGCTTGTCTCCACCGCAGTGTTA 2091
 Db 1 TGTCCATGTGATTAACCTGGGTACAGGGCCCATCTGCGCTCTCCACAGTGTGCA 60
 QY 2092 TCGTCTACTTGTGCTTGTGTTATGGATTTGGTCCATCCCAACATTTATGTGAG 2151
 Db 61 TCATCTACTTGTGCTTGTGTTATGGATTTGGTCCATCCCAACATTTATGTGAG 120
 QY 2152 AGATCTTCCCAACAGGGTGTGCGCTCTGTATTCGATTTGGCTTACATTTCTGA 2211
 Db 121 AGATCTTCCCAACAGGGTGTGCGCTCTGTATTCGATTTGGCTTACATTTCTGA 180
 QY 2212 TCGGATATGATGCTGCTACCTACAGCCTTCCGTGATGCTGATGCTATTTGAGTGGCGG 2271
 Db 181 TTGGAGATCATCTTTCACCTACAGCCTTCCGTGATGCTGATGCTATTTGAGTGGCGG 240
 QY 2272 GTGTTTTCAGATATATGACAGTGTGATGCTGATTTCTTTGTGTTCTTCTTAAAG 2331
 Db 241 GTGTTTTCAGATATATGACAGTGTGATGCTGATTTCTTTGTGTTCTTCTTAAAG 300
 QY 2332 TCCGTAACAAGAGGGAGTCCCTTGAAGTATTAACGAATCTTTGAGTGTGCGA 2391
 Db 301 TTCTCGAACAAGAGGGAGTCCCTTGAAGTATTAACGAATCTTTGAGTGTGCGA 360
 QY 2392 AGCAAGCGGTGCAAAAGCCTAATTTCTTGGTACCTTTGTGTGCAAC-TATTGACACTGT 2450

Db 361 AGCAAGCGGTGCAAAAGCCTAATTTCTTGGTACCTTTGTGTGCAAC-TATTGACACTGT 420
 QY 2451 AAGTTAAGAACTTGAAGGGTTTCCACCAAGAGCTCGAGATTAATCTTTGATTTGTGA 2510
 Db 421 GAATTAATTAACCTTGAAGGATCTCCACCAAGAGCTCTGGG-ACCTTCATTTGTATA 477
 QY 2511 AATGTTAAGGAACGAATCTGCTCAGTCCCTCAACGGTAAAGAGTCCCTCAAT 2570
 Db 478 AATGTTAAGGAACGAATCTGCTCAGTCCCTCAACGGTAAAGAGTCCCTCAAT 536
 QY 2571 GCGAATAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2629
 Db 537 GCGAAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 596
 QY 2630 TGTATTATAGTCAAGCTATTTCAACGCTGTTGTTGCTAGAAATCTTTGAACAAAGAT 2688
 Db 597 TGTATTATAGTCAAGCTATTTCAACGCTGTTGTTGCTAGAAATCTTTGAACAAAGAT 652

RESULT 13
 BM329724 661 bp mRNA linear EST 04-JAN-2002
 LOCUS PIC1 42 B11.g1_A002 Pathogen-infected compatible 1 (PIC1) Sorghum
 bicolor cDNA, mRNA sequence.
 ACCESSION BM329724
 VERSION BM329724.1 GI:18068861
 KEYWORDS EST.
 SOURCE Sorghum bicolor (sorghum)
 ORGANISM Sorghum bicolor
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Sorghum.
 REFERENCE 1 (bases 1 to 661)
 Cordonnier-Pratt, M.-M., Gingle, A., Fang, G.C., Dean, R., Wing, R.,
 Sudman, M. and Pratt, L.H.
 An EST database from Sorghum: plants infected with a compatible
 pathogen
 JOURNAL Unpublished
 CONTACT Cordonnier-Pratt MM
 Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 TEL: 706 542 1860
 FAX: 706 583 0210
 EMAIL: mmpratt@uga.edu
 SEQUENCES have been trimmed to exclude PolyA, vector, and regions
 below Phred quality 16. The threshold for highest quality sequence
 is 20. Three-prime sequences, which are obtained with PolyMix or
 T7 sequencing primer, are presented as the reverse complement.
 SEQ PRIMER: T7
 High quality sequence start: 57
 High quality sequence stop: 661
 POLYA=yes.

FEATURES
 SOURCE 1..661
 Location/Qualifiers
 /organism="Sorghum bicolor"
 /mol_type="mRNA"
 /cultivar="Btx623"
 /db_xref="taxon:4558"
 /tissue_type="Leaves"
 /dev_stage="4-week-old seedlings infected with
 Colletotrichum graminicola"
 /clone_lib="Pathogen-infected compatible 1 (PIC1)"
 /note="Vector: Bluescript II SK(-) from Lambda Zap II;
 Site 1: XhoI; Site 2: EcoRI; Four-week-old sorghum
 seedlings were sprayed with spore suspension prepared from
 3-week-old PM421, a sorghum isolate of the antherose
 pathogen Colletotrichum graminicola. Inoculated plants
 were kept in a 25 C dark growth chamber with 100% relative
 humidity for 24 hr, followed by 12/12 hr of light/dark
 cycle at 25 C with 90% relative humidity for another 24
 hr. All leaves were harvested and quick frozen with liquid

nitrogen and stored in a -80 C freezer. The library was made from poly-A RNA in the cloning vector lambda ZAP II. Clones to be sequenced were prepared by mass excision. WARNING: While most or all ESTs are expected to derive from the host plant, no effort was made to eliminate ESTs deriving from the pathogen."

BASE COUNT 164 a 138 c 144 g 215 t

Query Match 16.7%; Score 465; DB 12; Length 661;
Best Local Similarity 87.0%; Pred. No. 6.9e-46;
Matches 582; Conservative 0; Mismatches 75; Indels 12; Gaps 6;

QY 2080 CCGTACGATGTTACCTTACCTTCTGCTGCTTATGAGATTTGGTCCATCCCAACA 2139
DB 1 CAGTACGATGTTACCTTACCTTCTGCTGCTTATGAGATTTGGTCCATCCCAACA 60
QY 2140 TTTTATGTCAGAGATCTTCCACACGAGGTTGTCCTCTGTAATGTCATTTGCTT 2199
DB 61 TTTTATGTCAGAGATCTTCCACACGAGGTTGTCCTCTGTAATGTCATTTGCTT 120
QY 2200 TTACATTCGATGTCAGAGATCTTCCACACGAGGTTGTCCTCTGTAATGTCATTTGCTT 2259
DB 121 TGACGTTTGGATGTCAGAGATCTTCCACACGAGGTTGTCCTCTGTAATGTCATTTGCTT 180
QY 2260 TTGACGTCGAGGTTGTCCTCTGTAATGTCATTTGCTT 2319
DB 181 TTGACGTCGAGGTTGTCCTCTGTAATGTCATTTGCTT 240
QY 2320 TCTTCTTAAAGTCTTCCACACGAGGTTGTCCTCTGTAATGTCATTTGCTT 2379
DB 241 TCTTCTTAAAGTCTTCCACACGAGGTTGTCCTCTGTAATGTCATTTGCTT 300
QY 2380 CAGTTGGTCGAGGTTGTCCTCTGTAATGTCATTTGCTT 2439
DB 301 CAGTTGGTCGAGGTTGTCCTCTGTAATGTCATTTGCTT 360
QY 2440 -TATTGACATGTCAGAGATCTTCCACACGAGGTTGTCCTCTGTAATGTCATTTGCTT 2498
DB 361 TTTTTCGATGTCAGAGATCTTCCACACGAGGTTGTCCTCTGTAATGTCATTTGCTT 417
QY 2499 TGGATTTGTCAGAGATCTTCCACACGAGGTTGTCCTCTGTAATGTCATTTGCTT 2558
DB 418 TGGATTTGTCAGAGATCTTCCACACGAGGTTGTCCTCTGTAATGTCATTTGCTT 476
QY 2559 GAGTTCCTTCAATGTCAGAGATCTTCCACACGAGGTTGTCCTCTGTAATGTCATTTGCTT 2617
DB 477 GAGTTCCTTCAATGTCAGAGATCTTCCACACGAGGTTGTCCTCTGTAATGTCATTTGCTT 536
QY 2618 ACCATTTGTCAGAGATCTTCCACACGAGGTTGTCCTCTGTAATGTCATTTGCTT 2677
DB 537 ACCATTTGTCAGAGATCTTCCACACGAGGTTGTCCTCTGTAATGTCATTTGCTT 593
QY 2678 AGAACAAGATGTCAGAGATCTTCCACACGAGGTTGTCCTCTGTAATGTCATTTGCTT 2737
DB 594 AGAACAAGATGTCAGAGATCTTCCACACGAGGTTGTCCTCTGTAATGTCATTTGCTT 650
QY 2738 AATATGCTT 2746
DB 651 AATATGCTT 659

RESULT 14
CC004220 471 bp DNA linear GSS 31-MAR-2003
DEFINITION PUB0507D ZM 0.6.1.0 KB Zea mays genomic clone ZM05199103,
ACCESSION CC004220
VERSION CC004220.1 GI:29382780
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 471)
REFERENCE
AUTHORS Whitehead, C.A., Quackenbush, J., Van Aken, S., Uterback, T., Resnick,
A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, V. and Bennetzen, J.
TITLE
JOURNAL Maize Genomics Consortium
COMMENT Unpublished
CONTACT: Cathy Whitehead
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitehead@tigr.org
Seq primer: TP
Class: sheared ends.

FEATURES

source

1. 471
/organism="Zea mays"
/mol type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM05199103"
/clone_11b="ZM 0.6.1.0 KB"
/note="Vector: pCR4-TOP0; Site 1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"
BASE COUNT 115 a 83 c 155 g 118 t

Query Match 16.7%; Score 464.6; DB 29; Length 471;
Best Local Similarity 99.2%; Pred. No. 9.3e-46;
Matches 467; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1309 AGGCAACGTCGTCAGAGAGATGTCACCATGTCACCGTGAAGTGTGGA 1368
DB 1 AGGCAACGTCGTCAGAGAGATGTCACCATGTCACCGTGAAGTGTGGA 60
QY 1369 GCATGAGAGAGATGTCACCATGTCACCGTGAAGTGTGGA 1428
DB 61 GCATGAGAGAGATGTCACCATGTCACCGTGAAGTGTGGA 120
QY 1429 GTGGGAGATGTCACCATGTCACCGTGAAGTGTGGA 1488
DB 121 GTGGGAGATGTCACCATGTCACCGTGAAGTGTGGA 180
QY 1489 AAGGTCGTCAGAGAGATGTCACCATGTCACCGTGAAGTGTGGA 1548
DB 181 AAGGTCGTCAGAGAGATGTCACCATGTCACCGTGAAGTGTGGA 240
QY 1549 CAATGTCGTCACCATGTCACCATGTCACCGTGAAGTGTGGA 1608
DB 241 CAATGTCGTCACCATGTCACCATGTCACCGTGAAGTGTGGA 300
QY 1609 CTGCTTATGTCAGAGATGTCACCATGTCACCGTGAAGTGTGGA 1668
DB 301 CTGCTTATGTCAGAGATGTCACCATGTCACCGTGAAGTGTGGA 360
QY 1669 ATGTCGTCAGAGATGTCACCATGTCACCATGTCACCGTGAAGTGTGGA 1728
DB 361 ATGTCGTCAGAGATGTCACCATGTCACCATGTCACCGTGAAGTGTGGA 420
QY 1729 TTGAACCTGAGAGAGATGTCACCATGTCACCATGTCACCGTGAAGTGTGGA 1779
DB 421 TTGAACCTGAGAGAGATGTCACCATGTCACCATGTCACCGTGAAGTGTGGA 471

RESULT 15
CA502490 703 bp mRNA linear EST 14-NOV-2002
LOCUS CA502490
DEFINITION WHE4048_B04_D082T wheat meiotic anther cDNA library Triticum
ACCESSION CA502490
VERSION CA502490.1 GI:24993450
KEYWORDS EST.

SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
 ; Triticeae; Triticum.
 1 (bases 1 to 703)
 REFERENCE Anderson, O.D., Chao, S., Crossman, C., Langridge, P., Lazo, G.R., Pham
 J., Rausch, C.J., Sutton, T., Woo, D., and Wilson, C.
 TITLE The structure and function of the expressed portion of the wheat
 genomes - Meiotic anther cDNA library
 JOURNAL Unpublished
 COMMENT Contact: Olin Anderson
 US Department of Agriculture, Agriculture Research Service, Pacific
 West Area, Western Regional Research Center
 800 Buchanan Street, Albany, CA 94710, USA
 Tel: 5105595773
 Fax: 5105595818
 Email: oanderan@w.usda.gov
 Sequences have been trimmed to remove vector sequence and low
 quality sequence with phred score less than 20
 Seq primer: T7 primer.
 Location/Qualifiers
 1..703
 /organism="Triticum aestivum"
 /mol_type="cDNA"
 /cultivar="Chinese Spring"
 /db_xref="taxon:4565"
 /clone="WHE4048_B04_D08"
 /tissue_type="Anther"
 /dev_stage="Meiotic stages pre-meiosis-metaphase I"
 /lab_host="E. coli DH10B"
 /clone_lib="Wheat meiotic anther cDNA library"
 /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; Plants
 were grown in a glasshouse. Anther meiotic stage was
 determined by removing anthers from individual primary
 florets. One anther was sacrificed for microscopic staging
 , and if determined to be between (and including) meiotic
 stages pre-meiosis and metaphase I, the remaining two
 anthers were collected and pooled for library
 construction. The tissue, total RNA, and poly(A) RNA were
 prepared, cDNA synthesised, and directionally ligated into
 pSPORT1 by Tim Sutton in the P Langridge Lab at the
 Department of Plant Science, University of Adelaide, Waite
 Campus, Australia. Average insert size 1.5Kb. Plasmid DNA
 preparations and DNA sequencing were performed in the OD
 Anderson lab (all other authors)."
 BASE COUNT 139 a 183 c 165 g 216 t
 ORIGIN
 Query Match 16.6%; Score 460; DB 14; Length 703;
 Best Local Similarity 81.6%; Pred. No. 2.6e-45;
 Matches 545; Conservative 0; Mismatches 120; Indels 3; Gaps 1;
 QY 1602 CATGCGCTGCTTGTAGTACAGTACGACCTTTCTCAAGAGGCTTGCTGACACGCG 1661
 DB 36 CAGCGAGCTGCTTGTGTAAGCCACTGCGCTTTACTCCAGAGATTGAGAGAGCGT 95
 QY 1662 ATGTCAATGCT---GCCATGCTTCAACCATCTGAGGTAGTGCCTCAAGGTTCAAGTTGG 1718
 DB 96 ATGGGGGCTGCTCCAGCCATGATTCATTCATGAGGAGCGCTCCAAAGTTCAATCTGG 155
 QY 1719 AAAGATTGTTGAACTGAGTGAAGCGCTGCTTTAGTGTGCTGTTGGAATTCAATC 1778
 DB 156 AAAGATTGTTGAACTGAGTGAAGCGCTGCTTTAGTGTGCTGTTGGAATTCAATC 215
 QY 1779 CTTCACAGTTGCTGGAATAAAGCGTCTGTAATATATACCCCAATCTTGAGCAA 1838
 DB 216 CTTCACAGTTGCTGGAATAAAGCGTCTGTAATATATACCCCAATCTTGAGCAA 275
 QY 1839 GCTGCTGAGGAGTATCTTTCACAAATTGCTTCAGCTCGGACATCAGATCTCTTG 1898
 DB 276 GCTGCTGAGGAGTATCTTTCACAAATTGCTTCAGCTCGGACATCAGATCTCTTG 335

QY 1899 ATCAGTTCTCTACACTACTACTAATGCTTCCCTGATGAGCTTTGCCAGTCTGTATG 1958
 DB 336 ATCAGTTCTCTACACTACTACTAATGCTTCCAGATGAGTGTAGCATGACTATG 395
 QY 1959 GATCTTCCGGAAGAGGTTTTCCTGCTAGGACCAATTCGAATCTTGATGACATCTCTA 2018
 DB 336 GATATATCTGGAAGAGGTTTTCCTGCTAGGACCAATTCGAATCTTGATGACATCTCTA 455
 QY 2019 GTTATCTGCTGTTGCTGCTCAATCTAATGATTTGGGTACATGACCATCTTGTCTCC 2078
 DB 456 ATTTGTTGGTGTGTCCAAATGTTATCAACTGAGTACGGGCGCCACGCTGTGCTCTCC 515
 QY 2079 ACCGCTAGTGTATGCTACTCTGCTGCTTGTATGAGGATTTGGTCCATCCCAAC 2138
 DB 516 ACAGTTACGCTATGCTTCTCTGCTGCTTGTGCAAGGCTTGGGCCGATCCCAAC 575
 QY 2139 ATTTATGTGAGAGATCTTTCACACAGGCTTGAGCTCTGTATTTGCCATTTGTCGC 2198
 DB 576 ATTCATGTGAGAGATTTTCCCAACGAGCCCGTGTCTGTGATGCTATTTGGCGCC 635
 QY 2199 TTTACATTCTGAGTGGAGATATCATGCTCACTTACAGCTTCTGTGATGCTGATGCT 2258
 DB 636 CTCACATCTGGAATTTGACATTAATGTTACTTACAGCTGCTGTGATGCTGATGCT 695
 QY 2259 ATTGACT 2266
 DB 696 ATTTGCT 703

Search completed: January 5, 2004, 15:47:47
 Job time : 566 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2004, 09:46:20 ; Search time 10059 Seconds
(without alignments)
11293.969 Million cell updates/sec.

Title: US-10-051-909-31

Perfect score: 2777
Sequence: 1 caccggggttagatccggagc.....aaaaaaaaaaaaaaaaaa 2777

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rtd:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
1	2200.2	79.2	2824	6 AR208565	AR208565 Sequence
2	1642	59.1	2665	8 AY165599	AY165599 Saccharum
3	1075.4	38.7	2614	8 HV0534445	AJ534445 Hordeum v
4	946.8	34.1	137140	2 AP005756	AP005756 Oryza sat
5	940.2	33.9	2601	6 AR208568	AR208568 Sequence
6	932.4	33.6	2570	8 AY094465	AY094465 Arabidops
7	875.8	31.5	2190	6 AY506620	AY506620 Sequence
8	875.8	31.5	2190	8 ATH532570	AJ532570 Arabidops
9	806.2	29.0	2516	8 HV0534446	AJ534446 Hordeum v
10	700	25.2	2205	6 AX412656	AX412656 Sequence
11	700	25.2	2205	6 AX507559	AX507559 Sequence
12	691.2	24.9	1487	6 AR208571	AR208571 Sequence
13	672.2	24.2	2426	6 AR208571	AR208571 A thaliana
14	646.2	23.3	2190	8 ATH532571	AJ532571 Arabidops
15	637.8	23.0	135406	8 AC136843	AC136843 Oryza sat
16	599.2	21.6	86710	8 ATE23812	AL022604 Arabidops
17	599.2	21.6	197859	8 ATCHRIV83	AL161587 Arabidops
18	594.6	21.4	1692	6 AR208569	AR208569 Sequence
19	585.4	21.1	142114	8 AC072366	AC072366 Oryza sat
20	585.4	21.1	300957	8 AE017116	AE017116 Oryza sat
21	518.2	18.7	100693	2 AC121239	AC121239 Medicago
22	402.8	14.5	103192	8 AC007369	AC007369 Arabidops
23	402.8	14.5	143879	8 AC069251	AC069251 Genomic s
24	394.4	14.2	1009	6 AR208572	AR208572 Sequence
25	386.6	13.9	94349	8 ATE26013	AL133452 Arabidops
26	318.8	11.5	120538	2 AP004082	AP004082 Oryza sat
27	315.8	11.4	114918	8 AP004945	AP004945 Lotus jap
28	311.8	11.2	87592	8 AC099739	AC099739 Oryza sat
29	311.8	11.2	127098	8 AC144426	AC144426 Oryza sat
30	311.8	11.2	154128	8 AP000615	AP000615 Oryza sat
31	276.2	9.9	86162	2 AC126786	AC126786 Medicago
32	276.2	9.9	122107	2 AC144482	AC144482 Medicago
33	249	9.0	145365	2 AC127019	AC127019 Medicago
34	227	8.2	870	6 AR208567	AR208567 Sequence
35	225.2	8.1	70311	8 AF128457	AF128457 Oryza sat
36	220.4	7.9	77605	8 AF119222	AF119222 Oryza sat
37	218.8	7.9	142852	8 AF161269	AF161269 Oryza sat
38	194.4	7.0	108553	8 AF527809	AF527809 Sorghum b
39	185	6.7	106246	8 AF528565	AF528565 Zea mays
40	185	6.7	131457	2 AC144718	AC144718 Zea mays
41	185	6.7	346296	6 AF090447	AF090447 Zea mays
42	143.4	5.2	155	6 AR249238	AR249238 Sequence
43	123.2	4.4	510	6 AR208570	AR208570 Sequence
44	120.4	4.3	137568	8 AC093178	AC093178 Oryza sat
45	120.4	4.3	302228	8 AE017070	AE017070 Oryza sat

ALIGNMENTS

RESULT 1
LOCUS AR208565
DEFINITION Sequence 1 from patent US 6383776.
ACCESSION AR208565
VERSION AR208565.1 GI:21509752
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2824)
AUTHORS Allen,S.M., Hltz,W.D., Kinney,A.J. and Tingey,S.V.
TITLE Plant sugar transport proteins
JOURNAL Patent: US 6383776-A 1 07-MAY-2002;
FEATURES Location/Qualifiers

source 1..2824
BASE COUNT 644 a 649 c 745 g 775 t 11 others
ORIGIN
Query Match 79.2%; Score 2200.2; EB 6; Length 2824;
Best Local Similarity 90.2%; Pred. No. 0;
Matches 2372; Conservative 0; Mismatches 253; Indels 6; Gaps 2;
133 CTTACAGAGCCGGTGAACCTCGACGATCTTGAGAGACAAGATGTGGGGGCTGTTC 192
194 CTTGGCCCGCCGTCGGAGTGGGGCGGTAGATTTCCGGCGGCATATGGGGGGCCGCTGA 253
193 TTGTGCGCATAGTGCCTCCATCCGCAATCTATTGCAAGGGGGGAGCAATGCAACATCG 252
254 TGTGCGCATCGCGGCTCTTATCGGCACTTGCTCAGGGCTGGGACAAATGCAATTTG 313
253 CAGCTGCTGTCTGTATATTAAGAAGAAATTTCAATGCAAAATAGCCCACTGTGAGG 312
314 CTGAGCCGCTCTGTACATAAAGAGAAATTCACCTGCAAGCAAGCCCTGTATCGAAG 373
313 GACTAATTTGTGCAATGTCACTTATCGCGCCACCATGTTACTTATCTCCGGCCAT 372
374 GCTCATCGTGGCCATGTTCTCATTTGGGGCAACAGTCATCAACAATCTCCGGGGCCAA 433
373 TATCAGATCTCATTTGGCCGAGCCCTATGCTTATCTCTCTGCAATTTGTAATCTTTCA 432
434 GGGCTGACCTGGTGTGTAGAGAGCCCATGCTGGTCCCTGCTCTTACTTACCTTGTCA 493
433 GCGGCTCATCATGCTATGCTCTCTAAATGTCTATGCTCTGTGTGGAGAGCTTCGTAG 492
494 GTGGCTGTGATGCTTTTGGGGCCCAATTTGTATCATTTGCTCTCTCGAAGGCTCATTTG 553
493 ATGATTTTGTATTTGGCTTGGCTGTACGCTTGTGCTTGTGATTTTCAAAATATAGCC 552
554 ATGGGTGATATCGTTTGGGGGTCAACCTGTGCTCTCTCAATCCGAAATCGAC 613
553 C---TTGGAGATTAGAGTTTGTGATATCACTACCAATTCAGTGG--ATCAGAG 606
614 CGCACAGAAATTTTGGGGCTGNTGAAACAAGTCCGCAATTCATTTGGGGTCAAGNGAG 673
607 GAATTTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 666
674 GGAATTTCT 733
667 GAATTTATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 726
734 GGCATATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 793
727 ATCTTCTGATATCTTCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 786
794 ACTTGCCTGAAATCAACAAGTGGCTTGTNAGCAAAAGAAAGATGCGAGGGGAAGAG 853
787 TGTTCACAAAGTTACGGGGGAAAGAGATGCTCAGAGTGAATTTGCCCTTCTTCGAG 846
854 TGTTCACAAAGCTGCGGGGAAAGAGATGCTCAGAGGAGAGAGGCTCTTCTAGTTGAG 913
847 GGTTCAGATTTGAGAGAGACACTTTCATTTGAAGATCAATCTTGAACCTGCGACGAG 906
914 GTTTGGGGGTGGTAAAGATACAGATTTTNNAGATCATTTGAGACTGCGACGAGG 973
907 CAGCGGATGATCTTTTATCTAGCGGTGATTAAGAAACAATCACTTTATGGGCTTGAAG 966
974 CAGCGGATGATCTTTTATCTAGCGGTGATTAAGAAACAATCACTTTATGGGCTTGAAG 1033
967 AAGGCAAGTCAATGATGCTGAGACCTTCTAAGAGAGCCCATCATGCTGGAAGTGTGCTT 1026
1034 AAGGCAAGTCAATGATGCTGAGACCTTCTAAGAGAGCCCATCATGCTTGAAGTGTGCTT 1093
1027 CTTCTGATCTCTGATGAGAGATGAGAACAGATGATACCCCTTATGATTCGATTTG 1086
1094 CTTCTGATCTCTGATGAGAGATGAGAACAGATGATACCCCTTATGATTCGATTTG 1153

1087 TGACATTTTGTAGTGTCCATGAGATATGCTCAAGCTGAGAGATGATGAGAGCA 1146
1154 TGACATTTTGTAGTGTCCATGAGATATGCTCAAGCTGAGAGATGATGAGAGCA 1213
1147 CATTGTTTCCAACTTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1206
1214 CATTGTTTCCAACTTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1273
1207 AGTGGATGAAGAAATCTTATGAGATGACGAGAGATGATGATGATGATGATGATGATGATG 1266
1274 AGTGGATGAAGAAATCTTATGAGATGACGAGAGATGATGATGATGATGATGATGATGATG 1333
1267 GTGACTATGAGAGAAATCTTATGAGATGACGAGAGATGATGATGATGATGATGATGATGATG 1326
1334 GTGACTATGAGAGAAATCTTATGAGATGACGAGAGATGATGATGATGATGATGATGATGATG 1393
1327 GGAAGACATTTGTCACCATGCTGACCGTGAAGTGTCTTTGAGATGAGAGAGAGAGAGAG 1386
1394 GGAAGACATTTGTCACCATGCTGACCGTGAAGTGTCTTTGAGATGAGAGAGAGAGAGAG 1453
1387 TCTTAGGGAGGGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1446
1454 TCTTAGGGAGGGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1513
1447 CTTGGAATGAGTCAAGAGAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATG 1506
1514 CTTGGAATGAGTCAAGAGAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATG 1573
1507 TCTACTTGACCAAGAGAGAGTCTGCTGCTCAAGAGAGAGTCAATTTGTTCACTTCCG 1566
1574 TCTACTTGACCAAGAGAGAGTCTGCTGCTCAAGAGAGAGTCAATTTGTTCACTTCCG 1633
1567 GTGTCGAGATGTTTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1626
1634 GTGTCGAGATGTTTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1693
1627 CAGCACTTTTCTCAAGAGAGTCTGCTGCTCAAGAGAGTCAATTTGTTCACTTCCG 1686
1684 CAGCACTTTTCTCAAGAGAGTCTGCTGCTCAAGAGAGTCAATTTGTTCACTTCCG 1753
1687 CATCTGAGTACGCTGCAAGAGAGTCTGCTGCTCAAGAGAGTCAATTTGTTCACTTCCG 1746
1754 CATCTGAGTACGCTGCAAGAGAGTCTGCTGCTCAAGAGAGTCAATTTGTTCACTTCCG 1813
1747 GTGTCGAGATGTTTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1806
1814 GTGTCGAGATGTTTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1873
1807 TCTGATCTATACCCCAAAATCTTGAAGAGAGTGTGAGAGATTTCTTTCCAAAT 1866
1874 TCTGATCTATACCCCAAAATCTTGAAGAGAGTGTGAGAGATTTCTTTCCAAAT 1933
1867 TTTGCTCAGCTGCGCATCAGATCATCTTGAATCACTTCTCACTACCTTAATG 1926
1934 TTTGCTCAGCTGCGCATCAGATCATCTTGAATCACTTCTCACTACCTTAATG 1993
1927 TTTCTGATGATGCTTTGCCATGCTGCTTATGATCTTTCCGAGAGAGATTTTCTGCTC 1986
1994 TTTCTGATGATGCTTTGCCATGCTGCTTATGATCTTTCCGAGAGAGATTTTCTGCTC 2053
1987 TAGGACAAATTCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2046
2054 TAGGACAAATTCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2113
2047 ATTTGGATCACTAGCCCATGCTTGTCTTCCACGCTGATGATGATGATGATGATGATGATG 2106
2114 ATTTGGATCACTAGCCCATGCTTGTCTTCCACGCTGATGATGATGATGATGATGATGATG 2173
2107 GCTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2166
2174 GCTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2233
2167 GGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2226

327 CCTCTTCAGTTCGTACTTTGTTAGACGCTCATTATGCTATGCTGCTCCCGAATGTCTA 386
QY TGTCTCTGTTGGGACGCTTCTGTAGATGATTTGGTATTTGGCTTGGCTGTCACTTGT 526
Db TGTACTACTGCTGGGACGCTTATTAATGATTCGCTGTCCGCTTGGCTGTCACTTGT 446
QY GCCTTTGATTTTGAATAAGTACCCCTTCGGAATTTAGAGTTTGGCTGTAATACCTACC 586
Db TCTCTGTACATTTTCAAGAACGCTCTCCAGAGAAAGAGTTTGGCTGTAATACCTGCC 506
QY ACAATTCAGTGTAGAGAGAGATGTTCTGTCTCATCTGATGATGATTTGGAGTGTCT 646
Db ACAGTTTCACTGGGTGAGGGGGGATTTTTTGTGTGATCTGATGATGATGATGATGAT 566
QY GTGCGCATCACCCTTGGAGATTTATGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 706
Db GCTGCGCATCACCCTTGGAGATTTATGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 626
QY CTTTGGTTTGAATTTTATCTTCTGTAATCTCCAGATGCTCCAGATGCTCCGTTAGCAAGTGT 766
Db CTTTGGTTTGAATTTTATCTTCTGTAATCTCCAGATGCTCCAGATGCTCCGTTAGCAAGTGT 686
QY GATGCGAGAGGCAAAAAAGTGTGCAAAAGTTACGAGGGAAGAGAGATGCTCAGTGA 826
Db AATGCGAGAGGCAAAAAAGTGTGCAAAAGTTACGAGGGAAGAGATGCTCAGGCGA 746
QY ATGTCTCTCTCTCTCGAAGGTTGAGGTTGAGAGAGACACTTCCATTGAAAGATCAAT 886
Db AATGCTCTCTCTCTGTTGAAAGGTTGAGAGTCCGAGAGGATCTTCAATTTGAAAGATCAAT 806
QY CATGAGACCTGCGACCGAGGCGAGATGATCTTGTACTGACGGTGTGATGAGAGCAAT 946
Db CATGAGACCTGCGACCGAGGCGAGATGATCTTGTACTGACGGTGTGATGAGAGCAAT 866
QY CACACTTATGAGGCTGAGAGAGGCGAGTCAATGATGATGCTCGACTTGAAGGAGCCAT 1006
Db AATACTATTTGGGCTGAGAGGCGCATCANTGATGATGCTCGACTTGAAGGAGCCAT 926
QY CATGCTTGAAGTGTCTTCTCTTGTGATCTTGTGATGAGAGATGATGATGATGATGAT 1066
Db CATGCTTGAAGTGTCTTCTCTTGTGATCTTGTGATGAGAGATGATGATGATGATGAT 986
QY ACCCTTATGATCCGATTTGATGACATTTTGTGATGATGATGATGATGATGATGATGAT 1126
Db ACCCTTATGATCCGATTTGATGACATTTTGTGATGATGATGATGATGATGATGATGAT 1046
QY TGGAGAGATGAGAGACATTTGTTTCAAACTTTGAGATGATGATGATGATGATGATGAT 1186
Db TGGAGAGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1106
QY TCGAGATGAGAGATGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 1246
Db TCGAGATGAGAGATGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 1166
QY CGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1306
Db CGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1226
QY GCAGGCAACAGGCTGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1366
Db GCAGGCAACAGGCTGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1286
QY TGGTGGCGGATGAG 1406
Db TGGTGGCGGATGAG 1406
QY GAG 1546
Db GAG 1466

1547 CTCAATTTCTTCACTTCCCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1606
QY ATCTGTGTCTTCACTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1526
Db ATCTGTGTCTTCACTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1526
QY TGTCTCTTGT 1666
Db TGTCTCTTGT 1586
QY AGATGT 1726
Db AGATGT 1646
QY GTTTGAACCTGT 1786
Db GTTTGAACCTGT 1706
QY GTTTGT 1846
Db GTTTGT 1766
QY GGCAGTTATTTCTTCAAAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1906
Db GGCAGTTATTTCTTCAAAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1826
QY TGTCTACTACTTACTTATGCTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1966
Db TGTCTACTACTTACTTATGCTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1886
QY CGAAG 2026
Db CGAAG 1946
QY GGT 2086
Db GGT 2006
QY TGTATGT 2146
Db TGTATGT 2066
QY TGTCTACTACTTACTTATGCTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2206
Db TGTCTACTACTTACTTATGCTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2126
QY TGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2266
Db TGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2186
QY CTGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2326
Db CTGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2246
QY TGAAG 2386
Db TGAAG 2306
QY TGAAG 2446
Db TGAAG 2366
QY TGTGAAG 2506
Db TGTGAAG 2426
QY TGTGAAG 2566
Db TGTGAAG 2486
QY TGTGAAG 2615
Db TGTGAAG 2533

QY	Db	QY	Db	BASE COUNT	ORIGIN
2616	TTACCTATTGTACGATTTATTAAGTCAAGTATTCACGCTGGTGGTGAACAAATCT	2616	TTACCTATTGTACGATTTATTAAGTCAAGTATTCACGCTGGTGGTGAACAAATCT	563	38.7%; Score 1075.4; DB 8; Length 2614; Match Local Similarity 69.2%; Pred. No. 6.6e-272; Matches 1549; Conservative 0; Mismatches 666; Indels 24; Gaps 5;
2534	GTACTTATTTACTGATTATGATGATGTTGCTGTTATATAGTCTGCTTGTGGTGAAGATGT	2534	GTACTTATTTACTGATTATGATGATGTTGCTGTTATATAGTCTGCTTGTGGTGAAGATGT	586	755 g 708 c 2 others
2676	TT 2677	2676	TT 2677		
2594	TT 2595	2594	TT 2595		
RESULT 3					
HVU534445		2614 bp	2614 bp		
LOCUS		linear	linear		
DEFINITION		Hordeum vulgare mRNA for hexose transporter (stpl gene).	Hordeum vulgare mRNA for hexose transporter (stpl gene).		
ACCESSION					
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
SOURCE					

OY	165	TTGAGAGACAAGATGTGGGGGGCTGTTCTTGCCGCAATAGCGCTCCATGCGCAACTTA	224
Db	74	TCGGCGGCATGTGTGTGGGGCGCGCTGTGTGTGCCTATAGCGGCTCCATGCGCAACTG	133
OY	225	TTGCGAGGGGTGGGCAATATGCCAATCCATGCAAGCTGTCTTCTGTAATATAAGAAATTT	284
Db	134	CTGCAAGGGGTGGGTAATATGCGAACCATATGCAAGTGTCCCTCTGTACATAAAGAAAGATT	193
OY	285	CAATTGGAATAATAGCCCATCTGTGAGGGGACTAATGTGTCAATGTCACTTAATGGGGCC	344
Db	194	AGCTGGAAGACCAAGCCCTTGATTCAGAGGGCTCAATGTGCGCATGTGTCTATCGAGCG	253
OY	345	ACCATCTTATCTAATCTCTCCGGGCAATTAATACAGTCCGAATGGCCGACGCCCTATGCTT	404
Db	254	ACGGTTATACAGACGTTCTCCGGGGCGGTGGCTGAGCGCTGTGGTATAGCGGCCCTGTCTG	313
OY	405	ATTCTCTCTTCAATTTCTGTACTTCTTGAGGGGCTCATATGCTATATGTCTCTAATATGC	464
Db	314	ATCGCCCTGCTGTCCCTCTACTTGTGTCAAGTGGCTGTGTATGCTGTGGGGGCCCAAGTGC	373
OY	465	TATGCTCTGTGTGGAGCGCTTGATGATGATGATTTGGTATTTGGCTTGGCTGTACGCTT	524
Db	374	TATGTGTGTCTTTGGCGAAGCTCATGACGGGTGGGTATCGGTTTGGCAATACCTT	433
OY	525	GTGCTTTTGTACATTTGAGAAATAGCCCTTTCGGAGATTAAGAGTTTGTCTGAATATACTA	584
Db	434	GTCCCTCTGTACATTTGCGAGACCGCCCGCACTGACATTAAGAGGGCTGTGAACACGCTG	493
OY	585	CCAAATATTCATGTGATGAGAGAAATGTTCTGTGCATATGCATGATGATGTTGGATATGCC	644
Db	494	CCGAGTTTCATGTGGGTGAGAGGGATATCTCTTCTTACCTGCATGATGGTTTACCATGTGC	553
OY	645	CTGTGCGCATACCCCGATTTGGAGAAATATGCTGTGGTGTGCTCGGATACCTTCAATGTTC	704
Db	554	CTCATGTGCGAGCCTGACTGGAGAAATCATGTGTTGGGATTTTGTGATCCGTCGCTATG	613
OY	705	TTCTTTGGTTGACAAATATTTTATCTTCTGTGAATCTCCAAATGCTCTGTAGCAAAAGT	764
Db	614	TATTTTCAATTTGACTGTCTTTTATTTTTCGCCAGATCGCCCAAGATGGCTGTGTAGCAAGGA	673
OY	765	CGAATGCGAGAGCGCAAAAAGTGTTCAAAAAGTTACGGGGGAAAGCAATGTCTCAGT	824
Db	674	AGAAATGCTGAGGCCAAGCGAAGTATTCGAAAGACTGCGGGGAGAGGAATGTCTTCAGGA	733
OY	825	GAATTTGTCCTTCTTCTTCGAAGGGTGTGAAGTTGAGAGAGACACTTCATTTGAAGATAC	884
Db	734	GAGATGCCCCCTTTGTTGTAAGGATTTGGGTGTTGAAAAGACACAACTTTTGAAGAAATAC	793
OY	885	ATCATTTGACCTGTGCACCGAGGCGACGCCAGATTTTGTATCTGACGGTGTATTAAGAACAA	944
Db	794	ATAATTTGACTCTGTATGTATGTAGCTGGCTGTATGACGAGCTGGCTGCT---CCAGATCAAGAGAG	850
OY	945	ATCACACTTTATGGGCTCTGAAGAAGCGCAGTCAATGATTTGCTCGACCTTCTAAG-----	998
Db	851	TTGAAATCTTATATGAGCTGAAGAAGGGATATCTTGATATGCGCGCTCTGTTAGGSSGTGCG	910
OY	999	GGACCCATCATGCTTGGAAATGTGTCTTCTTTCATCTGTCAATGGAGCATGTGAAC	1058
Db	911	GGCCAAAGTGCATTTGAAAGCGCTTTTGGGCTCAATGTCTGTCAATGGAGATATGTATGT	970
OY	1059	CAGAGTGAACCCCTTATGATCCGATTTGTGCACTTTTGGTATGTATGTCATGAAGATATG	1118
Db	971	CAGGTTAATATCTTGTGTGACCCCTTGTCACTCTTTCGGAAATGTTCATGAAAGATG	1030
OY	1119	CCTCAAGCTGAGGAATATGAGAGACAAATTGTTTCCAACTTTGGAAGTATGTCAGT	1178
Db	1031	CCTGAGTTAATATGGGAGACATGCGGAGACCTGTGTTTCTTAACTTTGACAGCATATTAGT	1090
OY	1179	GTCAACATATGCAATGCGCAAAAATGAGACATGGGATTAAGAGATCTTCAATAGGATGAC	1238
Db	1091	GTGGCTAAACGACGACAGGCTTAAAGCTGATCGGATCTGTAGA---GTCAATGGGAATGAT	1147

QY 1239 GAGAGTACGATCTGATGCTGAGAGGTGACTATGAGACATCTCCATGCCCCATG 1298
 DB 1148 GAAAGATTATGACATGATCATGATGCTGATGACATGAGATTAACCTCAATGCCGCTT 1207
 QY 1299 CTGTCAGGACGAGCAACAGGTGCGGAGGAGGAGCACTATGTCACCATGTCACCGTGA 1358
 DB 1208 ATTCTTCGTCAGACGACAGCGGTGAGGAGTAAAGAAATTCGTCACCTCATGAAAGCAT 1267
 QY 1359 AGTGTCTTGAAGCATGAGAAAGCAACCTCTTAAAGGAGGTGAGAGTGTGAGCAGC 1418
 DB 1268 ATGGGT-----GCTGTGAAAGATAGTACATGACAGGAGGAGGAGCAGTAAGCAGC 1318
 QY 1419 ACTGATTCGAGGAGGAGTGAAGTGTGTTGAAATGCTGAGAGAGAGAGTGAAGAT 1478
 DB 1319 ATGGCATTTGGTGGGGGTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1378
 QY 1479 GGTAGAAAGAGAGTGTTCCTTCAAAAGATCTACTTTCACCAAGAGGAGTCTGCTGCA 1538
 DB 1379 GGGCAAAAGAGAGTGTTCCTTCAAAAGATCTACTTTCACCAAGAGGAGTCTGCTGCA 1438
 QY 1539 AGAAGGGCTCAATGTTTCACTTCCGGTGTGAGATGTTTGAAGGTAGTGAAT 1598
 DB 1439 CGGAGAGGCTCTATATGCTATGCTAGCAGGAGGT--GATATTCCTCTGCTGCTGAT 1495
 QY 1599 GTACATGCTGCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1658
 DB 1496 ATCCAGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1555
 QY 1659 CGCATGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1718
 DB 1556 CAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1615
 QY 1719 AAGATTTGTTGAACTGAGTGAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1778
 DB 1616 GCAAGAACTATTTGAACCTGAGTGAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1675
 QY 1779 CTTCACAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1838
 DB 1676 CTGACGAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1735
 QY 1839 GCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1898
 DB 1736 GCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1795
 QY 1899 ATGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1958
 DB 1796 ATTAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1855
 QY 1959 GATCTTCCGAGAGAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2018
 DB 1856 GATATGTGAGAGAGAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1915
 QY 2019 GTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2078
 DB 1916 GCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1975
 QY 2079 ACCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2138
 DB 1976 ACGATCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2035
 QY 2139 ATTTATGAGAGAGATCTTTCACCAAGAGGTTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2198
 DB 2036 ATTTCTGCGGAGAGATTTTCCCACTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2095
 QY 2199 TTTACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2258
 DB 2096 CTGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2155
 QY 2259 ATTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2318
 DB 2156 ATCCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2215
 QY 2319 GTCTTCTTAAGGTCCCTGAGCAAGAGGAGTCCCTTGAAGTTTAAACGAATTCCTT 2378

DB 2216 GTTACATGAAGGCTCCCTGAGACAAAGGCAATGCCCTGAGGCTCATCCAGTTCTTC 2275
 QY 2379 GCAGTGTGCTGCGAAGCAAG 2397
 DB 2276 TCGGTGGGGGAGAGCAGG 2294
 RESULT 4
 AP005756/c 137140 bp DNA linear HTG 19-SEP-2002
 LOCUS Oryza sativa (japonica cultivar-group) chromosome 2 clone
 DEFINITION OSJNB0035N08, *** SEQUENCING IN PROGRESS ***
 ACCESSION AP005756
 VERSION AP005756
 KEYWORDS HTG; PHASE2.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE 1
 Sasaki, T., Matsumoto, T. and Katayose, Y.
 Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC
 clone: OSJNB0035N08
 Published Only in Database (2002)
 2 (bases 1 to 137140)
 Sasaki, T., Matsumoto, T. and Katayose, Y.
 Direct Submission
 Submitted (18-SEP-2002) Takuji Sasaki, National Institute of
 Agrobiological Sciences, Rice Genome Research Program, Kamondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail: tsasaki@nias.affrc.go.jp, URL: http://rtp.dna.affrc.go.jp/
 Tel: 81-298-38-7441, Fax: 81-298-38-7468)
 NOTE: It currently consists of 1 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces is believed
 to be correct as given, however the sizes of the gaps between them
 are based on estimates that have provided by the submitter. This
 sequence will be replaced by the finished sequence as soon as it is
 available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 FEATURES
 source
 1. 137140
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:3947"
 /chromosome="2"
 /clone="OSJNB0035N08"
 BASE COUNT 40995 a 27904 c 27671 g 40414 t 156 others
 ORIGIN
 Query Match 34.1%; Score 946.8; DB 2; Length 137140;
 Best Local Similarity 69.9%; Pred. No. 7.2e-228;
 Matches 1520; Conservative 0; Mismatches 407; Indels 249; Gaps 8;
 DB 86221 ATTGCGAGAGAGATGCTCTTCTTGTGAAGTTTGAAGTTGAGCTGACCTTCCA 86162
 QY 874 TTGAAGAGTACATGATGAGACTGCGACCGAGCGAGTATGTTACTGACGGTG 933
 DB 86161 TTGAAGAGTACATGATGAGACTGCGCTATGAGCGACGATGACATGTTGATGCG 86102
 QY 934 ATTAAGAAACAATGACATTTATGAGGCTGAGAGAGCGCATGATGCTGACCTT 993
 DB 86101 ATTAAGAAACAATGACATGATGAGGCTGAGAGAGCGCATGATGCTGACCTT 86042
 QY 994 CTAGGAGCCATGATGCTGAGAGTGTGCTTCTTGTGATCTGCTGAGGAGCATGG 1053

Db 86041 CCAAGGACCCAGCATCTTGAGAGTGTCTTCTTACATCTGTCATGCGACATGG 85982
Qy 1054 TGAACCAAGATGATCCCTTATGATCCGATTTGTGACATTTTGGTATGTCATGAGA 1113
Db 85981 TGAACCAAGATGATCCCTTATGATCCGATTTGTGACATTTTGGTATGTCATGAGA 85922
Qy 1114 ATATGCTCAAGCTGAGGAAAGTATGAGAGACATTTGTTCCAACTTTGGAAATATGT 1173
Db 85921 ACATGCTCATGCTGAGGAAAGCATGCGAGATGATGTTTCCCACTTTGGTATATGT 85862
Qy 1174 TCAGTGTACAGATATGACATGCGCAAAATGACAGTGGGATGAGAGAAATCTTCAAGG 1233
Db 85861 TTAGTGTACAGATATGACATGCGCAAAAGTGTATCAATGGAGATGAGAGAACTTCAAGG 85802
Qy 1234 ATGACGAGAGATGACATCTGTATGAGAGAGTGTATGAGAGCAATGTCATGACCC 1293
Db 85801 ATGATGAGAGATGACATCTGTATGAGAGAGTGTATGAGAGCAATGTCATGACCC 85742
Qy 1294 CATGCTGTCCAGGAGGCAACAGGTGCGGAGAGGAGAGCAATGTCATGACCC 1353
Db 85741 CGTTGCTGTCCAGGAGGCAACAGGTGCGGAGAGGAGAGCAATGTCATGACCC 85682
Qy 1354 GTGGAAGTCTTGTGAGCATGAGAGGCAACCTCTTGAAGGAGGAGGAGAGTGTGA 1413
Db 85681 GTGGAAGTCTTGTGAGCATGAGAGGCAACCTCTTGAAGAGGAGGAGGAGGAGTGA 85622
Qy 1414 GCAGCATGATATGCTGAGGAGAGTGGAGCTTGTGAGAAATGTCAGAGAGAGAGTGTG 1473
Db 85621 GCAGCATGAGCATGAGGAGAGTGGAGCTTGTGAGAAATGTCAGAGAGAGAGTGTG 85562
Qy 1474 AGAATGTAGAGAGAGAGAGTGTGTTCAAAAGAGTCTTGTGACACCAAGAGAGTGTG 1533
Db 85561 AGAATGTAGAGAGAGAGAGTGTGTTCAAAAGAGTCTTGTGACACCAAGAGAGTGTG 85502
Qy 1534 GCTCAAGAGAGGAGCTCAATGTTTCACTTCCGAGTGTGAGAGTGTGAGAGTGTG 1593
Db 85501 GATCAAGAGAGGAGGCTCAATGTTTCACTTCCGAGTGTGAGAGAGTGTGAGAGTGTG 85442
Qy 1594 AGTTGTATCATGCTGCTCTTGTATGATGATGATGATGATGATGATGATGATGATG 1653
Db 85441 AATTCATATCATGCTGCTCTTGTATGATGATGATGATGATGATGATGATGATGATG 85382
Qy 1654 AACCAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1713
Db 85381 AACCAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 85322
Qy 1714 GTTGAAGAGATTTGTTGAACCTGAGAGTGGAGGCTGCTGTTAGTGTGAGATTC 1773
Db 85321 GCTGGAAGATTTGTTGAACCTGAGAGTGGAGGCTGCTGTTAGTGTGAGATTC 85262
Qy 1774 AGATCTCTCA----- 1784
Db 85261 AAATCTCTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 85202
Qy 1785 ----- 1784
Db 85201 TTGACGATGTTAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 85142
Qy 1785 -----CAGTTGCTGAATAAAGGATGATGATGATGATGATGATGATGATGATG 1838
Db 85141 CTTGACGATTTGCTGATTAATGAGGATGATGATGATGATGATGATGATGATGATG 85082
Qy 1839 GCTGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1898
Db 85081 GCGGGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 85022
Qy 1899 ATCAGTTCTCTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1958
Db 85021 ATCAGTTCTCTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 84962
Qy 1959 GATCTTTCCGAGAG----- 1973
Db 84961 GACATCTCTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 84902

Qy 1974 ----- 1973
Db 84901 AATATACCAAGAGAGATGACCCCAAGTATGATTAATTAATGATGATGATGATGATG 84842
Qy 1974 ---AGTTTGTCTGCTAGGAGCAATTCATCTTGTATGATGATGATGATGATGATG 2030
Db 84841 TTAGATTTCTGCTTCTGAGGAGCAATTCATCTTGTATGATGATGATGATGATGATG 84782
Qy 2031 GTTTCATCTTAATGATTTGGGATGATGATGATGATGATGATGATGATGATGATGATG 2090
Db 84781 GTTTCATCTTAATGATTTGGGATGATGATGATGATGATGATGATGATGATGATGATG 84722
Qy 2091 ATGCTACTCTCTGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2150
Db 84721 ATCATCTACTCTCTGCTCTGATGATGATGATGATGATGATGATGATGATGATGATG 84662
Qy 2151 GAGATCTTTCCAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2210
Db 84661 GAGATCTTTCCAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 84602
Qy 2211 ATGGAATATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2270
Db 84601 ATGGAATATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 84542
Qy 2271 GGTGTTTTCAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2330
Db 84541 GGTGTTTTCAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 84482
Qy 2331 GTCCCTGAGCAAGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2390
Db 84481 GTCCCTGAGCAAGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 84422
Qy 2391 AAGCA---GCGGCTGCAAAAGCTTAATTTCTTTGTGAC---CTTTGTGCAACTATT 2443
Db 84421 AAGCAAAATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 84362
Qy 2444 GCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2503
Db 84361 ACATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 84304
Qy 2504 TTGATTAATGATTAAGGAGAGCAATCTGCTCATGCTCTCTCAAAACGATTAAGAGATC 2563
Db 84303 GT-----AATGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 84250
Qy 2564 CCTCAATGCAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2623
Db 84249 CATTAAGGAGAGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 84190
Qy 2624 TTGATCTGATTA-----TAGTCAGCTTAATGAGGCTGCTGCTGATGATGATGATGATG 2674
Db 84189 TGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 84130
Qy 2675 TTAGAACAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2734
Db 84129 TTAGAACAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 84072
Qy 2735 AAAATATGTTTCTCA 2750
Db 84071 TAAATTTCTGCTGCA 84056

RESULT 5
AR208568
LOCUS AR208568 2601 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 7 from patent US 6383776.
ACCESSION AR208568
VERSION AR208568.1 GI:21509755
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2601)

AUTHORS Allen, S.M., Hiltz, W.D., Kinney, A.J. and Tingey, S.V.
 TITLE Plant sugar transport proteins
 JOURNAL Patent: US 6383776-A 7 07-MAY-2002;
 FEATURES Location/Qualifiers
 source 1..2601
 /organism="unknown"
 BASE COUNT 648 a 543 c 644 g 766 t
 ORIGIN
 Query Match 33.9%; Score 940.2; DB 6; Length 2601;
 Best Local Similarity 66.0%; Pred. No. 2.9e-236;
 Matches 1476; Conservative 0; Mismatches 713; Indels 48; Gaps 6;

177 ATGTGGGGGCTGTCTTGTGCGCATATGCGCTCCATGCGCATTTATGCGGGGTG 236
 Db 175 ATGAAAGGCGCGTCTTGTGCTATGCGGCTCCATGCTATTTCTCCAGAGATG 234
 Qy 237 GACAAATGCCACCATGCGACGCTGTCTGTATATATAAGAGAAATTCATTTGCAAAAT 296
 Db 235 GATTAATGCTACCAATGCCGGGCTAATGTTACATTAGAAAGACCTTGCTTTGGGA-- 292
 Qy 297 GAGCCCACTGTGAGAGGACTAATGTGTCAATGTCACTTATCGGCGCCACCATGCTTACT 356
 Db 293 ---CACTATGAGAAAGCTGTGTGTGCGCATGTCCCTGATTGAGCAACGGTAATCAC 348
 Qy 357 ACATTTCCGGGCGCATATACAGACTGATGGCCACGCCCTATGCTTATTTCTCTTCA 416
 Db 349 ACATCTCTGTGCTCATATACGATGGCTGCGGTGCGGACCATATATATCTCATCT 408
 Qy 417 ATTCTGACTCTTTCAGCGGCTCATATGCTATGCTCTGCTATATGCTATGCTGCTG 476
 Db 409 GTGCTCTATTTCTTGGGTGTTGGTGTGATGCTGTGTGCTCCCAATGTATGTGTTGTC 468
 Qy 477 TTGGACGCTTGATGATGATTTGTGATATGCTGTGCTGTCACGCTGTGCTTTGATC 536
 Db 469 TTGGGAGGCTACTGATGATTTGGGATTTGGCCTTGCTGTCATCTTGCTCCGGCTAT 528
 Qy 537 ATTTCAGAAATAGCCCCCTTCGAGATTTAGAGTTTGTCTGAATACATACATTTCACT 596
 Db 529 AATATCTGAAACGGCGCTGTGAATTAAGGGGTGTTGAATACGCTTCTAGTTCACT 588
 Qy 597 GGATCAGAGAGATGTTCTGTCTATCTGATGATGTTGGATGCTCCGTGCGCATCA 656
 Db 589 GGCCTGTGAGAGATGTTTGTGCTATGATGTTTGGATGCTGATGCTGCTGCTGCTG 648
 Qy 657 CCCGATTGAGAAATTAAGCTGTGTGCTGCGCATACCTTCAATGTTCTTTGATTTG 716
 Db 649 CCTACTGAGGCTCATGCTTGGGGTCTGTCTATCTCTCTGTATTTTGCATTTG 708
 Qy 717 ACAATATTTTATCTTCTGATATCTCCAGATGCTTCGTTAGCAAGGTGAGATGCGAAG 776
 Db 709 ACCATTTTTTCTTCCCGAGTCTCTCGTGGCTGTGTCAGAAAGAGATGCTCAG 768
 Qy 777 GCAAAAGAGTGTGCAAAAGTTACGGGGGAAAGCATGTCTCAGGTAAATTTGCCCT 836
 Db 769 GCTAAGAGGTCTCCAAAGATTGGCGGAAGAGATGTCTCAGGCGAGATGGCATTTG 828
 Qy 837 CTTCGAGAGGTGAGGTGAGAGAGACATTCATTGAAGATACATCATTTGAGACT 896
 Db 829 CTGGTTGAAGTCTCGGAGATGGGGGTGATATCATGATGAGAGATATATTTGGCCCT 888
 Qy 897 GCGACCGAGGCGCGGATGATCTTGTACTGACGGTATAGAGAAATACATCTTAT 956
 Db 889 GCTGACGATGTGCTGATGTGATGAAACATGCAACAGAAAGATTAATTTGATTAAT 948
 Qy 957 GGGCTGAAAGAGCGCATGATGATGCTGACCTTTAAGAGCCCATGATGCTTGA 1016
 Db 949 GGAATCCAGAGCGCTTTCTTGTTATCAAAACCTGTCACTGACAGAGTTCTATTGG- 1007
 Qy 1017 AGTGTCTTCTTCTGATCTGTCAATGAGAGATGTGAACCAAGATGTACCCCTTATG 1076
 Db 1008 -----CCTTGCTCAACCATGAGATGATATCAACCAAGATGCCCCCTCATG 1056

Qy 1077 GATCCGATTGACATTTTGTAGTGTCCATGAGAAATAGCTCA-----AGCTGA 1130
 Db 1057 GATCTCTGTGACACTGTGTGTGACATTTATGAGAACTCCCGAACAAGAGAA 1116
 Qy 1131 GAAATGATGAGAGACATTTGTTCCAACTTTGGAAGTATGTTCACTGATCAAGATGAG 1190
 Db 1117 GGAAGCATGCAAGACTCTCTTTCCAAATTTTGGAGCATGTTTCAAGCATGCTGAGCG 1176
 Qy 1191 CATGCCAAAATAGACAGTGGATGAGAGAAATCTTCACTAGGATGACGAGATGCGCA 1250
 Db 1177 CATGTAAATTTGAACAATGGATGAGAGAAAGCTTACAAAGGAAACGTGAGACTATG 1236
 Qy 1251 TCTGATGTGAGAGAGTATGAGGACATCTCCATAGCCATGCTGTCCAGCGAG 1310
 Db 1237 TCAATGCAACCGGTGGGACTCCGATGATATTTGACATCTTTAATCTCACGCCAA 1296
 Qy 1311 GCAACAGGTGCGAAGGAGAGAAATGTGACATGTGACATGCTGCAAGTCTTTGAGC 1370
 Db 1297 ACACACAGCTTGAAAAAGACTTATACCTCTCTCTCCATGGCAATATCTTGCGAGC 1356
 Qy 1371 ATGAAAGCA--AACCTCTTAGGGGAGGTGAGATGTGTGACGACATCTGATATC 1427
 Db 1357 ATGAGCGCTCACAGTATCTCATGCAAGGCTCAGGTGAGCAAGGTGATGATATTT 1416
 Qy 1428 GTGGGGGATGCGAGCTTGTGGAATGTCTAGAGAAAGGATGGAATGTGAGAAG 1487
 Db 1417 GTGTGTGCTGCAACTGCGATGAAATGACTATTA--AGTGAAGATGAGAAACAA 1473
 Qy 1488 GAAGTGTGTTCAAAAGAGTCTACTTGACCAAGAGGAGTCTGCTGCTCAAGAGGGGC 1547
 Db 1474 CAAGAGGCTTAAAGATTTATTTACATGAGAGGAGATTTGTGATCTGCTGTGGA 1533
 Qy 1548 TCAATGTTTCACTCCCGGTGTGCGGAGTCTTTGAGGTAGTGAATTTGATCATGCT 1607
 Db 1534 TCAATGATGATTTCCCGTGAAGGCGAA-----TTGTCCAGGCT 1575
 Qy 1608 GCTGCTTAGTAATGCTGACGACTTTCTCAAGAGGCTTGTGGAACCAAGCATGCA 1667
 Db 1576 GCTGCTGTGTAACCAACCGCTCTTACTCAAGAGGCTTATGATGACACCCAGTT 1635
 Qy 1668 GATCTGCGCATGTTTCAACCATCTGAGGATGCTCCAAAGGTTGACGTTGAAAGATTTG 1727
 Db 1636 GGGCTGCAATGTTTCAACCATCTGAGACAGCTTCAAGGGCCAAAGTTGAAAGGCTTT 1695
 Qy 1728 TTGAACCTGAGTGAAGCGTGCCTGTATGCTGTGATGCTGATTTCAAGATCTTCAACG 1787
 Db 1696 CTGAACCAAGGGTTAAGCATGATGTTGTTGAGTGAATTAACAAATCTTCAGAG 1755
 Qy 1788 TTGCTGGAATAACGGGTCTGTACTATACCCCAAAATCTTGAGCAAGCTGTGTG 1847
 Db 1756 TTTTCAGGATTAATGGGTTCTATATTTACACACTTCAATCTTTGAGAGGCGGTGTT 1815
 Qy 1848 GCATTAATCTTCCAAATTTGTCTCAGCTCGGCATGAGCATCATCTTGATGATGTTCT 1907
 Db 1816 GAATGTTCTTTCAAGATATGACATTTGGCTCAAGTGGCATTTCTTATCAGTGCT 1875
 Qy 1908 CTCACATCTTACTAATGCTTCTTGCAATGAGCTTTGCAATGCTGCTTATGATCTTTCC 1967
 Db 1876 TTCAACAACCTTCTGATGCTCCCTGTATAGGCTTACGATGAGATGATGATTTCA 1935
 Qy 1968 GAAAGAGGTTTGTGCTAGAGCAAAATTCATCTTGATGAGATCTAGTATATCTG 2027
 Db 1936 GGCAGAAAGCGTGTCTACTATCAATCCCGTGTGATGTGTCACTATATTTTG 1995
 Qy 2028 GTTGTGTCATCTAATTTGATTTGGGTACATGACCATGCTTGTCTCCACCGTCACT 2087
 Db 1996 GTCAATGGAAGCTGTGTAATTTTGGCAATGTGCGCATGAGCAATCTCAACAGTATG 2055
 Qy 2088 GTTATGCTTACTTCTGCTGCTGTTATGGAATTTGGTCCATCTCCCAACATTTTATGT 2147
 Db 2056 GTTGTGTTATTTCTGCTGTTGTGATGGTATGAGCAAAATTCAAACATCTTTTGC 2115
 Qy 2148 GCAAGATCTTTCACACAGGTTCTGTGCTCTGTATGCAATTTGTGCTTTACATTC 2207

Db	2116	TCAGAGATTTTCCCACCTAAGGGTGGTGCCCTCGCATTTGCTATCTGTGCAATTAGTGTTC	2175
Qy	2208	TGGATTCGGAATATCATCTGTCACCTACAGCCTTCTCTGAAATGCTGAATGCTATTGGACTG	2267
Db	2176	TGGATTGAGACATCATCATCAACATACGCTGCTGCTGTGAATGCTCGGCTCTTTAGACTT	2235
Qy	2268	GCGGGTGTTTTCAAGCATATATGACAGCTGATGCTGATTTCTTTGTGTGTCCTT	2327
Db	2236	GGTGTGTATTTGGCATATTACGACAGTTGTTGTTTCATCTGATGAATATTTGTGTTTTTG	2295
Qy	2338	AAGGTCCTTGAGACAAAGGGATGCCCCCTTAGGTTATTACGAATCTTTGCAAGTGGT	2387
Db	2296	AAGGTTCCAGAAACAAAGGCGATGCCCTTTGAAGTCATCTGAATTTCTTTTCTGTGGA	2355
Qy	2388	GCGAAGCAGCGGCTGC 2404	
Db	2356	GCAAGACGAGCTGCTTC 2372	

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
6	AY094465	AY094465	Arabiidopsis thaliana	Arabiidopsis thaliana	Arabiidopsis thaliana (thale cress)	Arabiidopsis thaliana	Arabiidopsis thaliana
		2570 bp	Ar4g35300/P23E12_140	GI:20453188			
		mRNA					
		linear					
		PLN 05-MAY-2002					
			mRNA, complete cds.				

REFERENCE
AUTHORS
1 (pages 1 to 2570)
Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Bath, J.,

TITLE	JOURNAL	Unpublished
Arabidopsis cDNA clones	inocogis/n. and chncl, u.s.a.	

TITLE Direct Submission
JOURNAL Submitted (08-APR-2002) Salk Institute Genomic Analysis Laboratory

COMMENT

The Salk, Stanford, POEC (SSP) Consortium members carried out the sequencing and annotation of the RALV cDNAs: Shin, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Bahr, J., Bowser, L., Chang, B., Dale, J.R., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Torti, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Shinn, P. (SSP/Saik) and Seki, M. (RIKEN GSC) contributed equally to

this work. Shinozaki, K. (RIEN GSC) and Ecker, J.R. (SSP/salk).
 contributed equally to this work as PIs.
 Location/Qualifiers
 1. 2570
 FEATURES
 SOURCE

CDS

```

/note="putative sugar transporter protein"
/codon_start=1
/product="ATG35300/F23E12_140"
/protein_id="AAM9835.1"
/db_xref="GI:20453189"
/translation="MSGAVVAVIAAANGNLLCGMDNATIGAVVYIKKEPNTSEPSV
EGLVVAMSLIGATLITTCSSGVADMRLRRPMLLISSLIYYGSLWMLSPVYVLLLG
RLLDMFGVGLVTVLTPVPIYISTAPFELRGLNTLPQTSGGMFLSTCGMFGNLMPS
PMLMGVGLVPIBSLVFFLVFFLPSPMLVSGMGLSEKARVLOLRREDEVGSEM
ALVLEGISIGETTIEEYIIGPADEVETDHDIAVDKQIKYAGEBLISWAARVKVGG
STMSVLSRHGSTMSRQGSIDLPVTLFVGHSEHMDPTGMSRALPHFGFMEFGVAG
QPRHEDDENLVVEGEDGSDGMDHDSIEDLSPLIRPTTSMKMDPHAGLSTFTEF
RHGGVOVQAGQEGDSMGVCGMOWAKRTERBEDSOKEGFRYVILHOGFPGSR
GSVTSIPGSGTGEDFVQOASVATSOPLATYSKOLLKHTTIGPMVHSEITTKGSIMD
LHDSGVRVAVYGVGLOILOOFGSINLVVYTPLODAGCVTLISNMGSSSSASL
ISALTITVMLPALAVANRLMDLSGRRTLLTTPIILASLVIVTSLVHNNSTVAV
LSTSVLVLCFFVWVGSPAPNLLVCSLFFPRVANGICIAALTFWICDILVITSLP
LKSLGIAGVGMAIVCCISWVEFLFKVETPKGMPLEIVTEPFSVGARQEAANE

```

Query Match	33.6%;	Score 932.4;	DB 8;	Length 2570;
-------------	--------	--------------	-------	--------------

Matches 1474; Conservative 0; Mismatches 736; Indels 30; Gaps 7;

168 GAGGACAAGATGTCGGGGGCTTCTTGTGCGCCATAGTCGCCCTCCATCGGCAATCTATTG 227

[illegible]

A vertical ruler scale with markings from 2.20 to 2.27. The scale is oriented vertically, with the numbers 2.20, 2.21, 2.22, 2.23, 2.24, 2.25, 2.26, and 2.27 arranged from top to bottom. Each integer part of the scale (2.20 to 2.26) is subdivided into millimeters, with horizontal lines indicating each millimeter increment.

260 C A G G A T G G G A T A A C C G A A C T A T T G C A G S A G C T G T G T G T G T A C A T A A A A A A G G A G T T T A A T 319

288 TTGCAAAATGAGCCCACTGTGGAGGACTAATTGTCTCAATGTCACCTTATCGGCCACC 347

320 TTGAGAGTAATCCATCAGTGAAGGCTAATTGTGGCGATGTCACCTTATTGGTCTACT 379

348 ATCGTACTACATTCCTCCGGCCATTATCAGACTCGATTGGCCGACGCCCTATGCTTATT 407

[illegible]

[illegible]

440 IIGICLCMAIILCIIACIIIGIIGOIILCIIAGIAHIGIIMIGIILCCGAHIGIIM 455

468 GTCTGCTGTTGGCACCCTTCGTAGATGGAATTGGTATTGGCTTGGCTGTCAACGCTTGTG 527

500 GTGTTGCTCTTAGGAAGTTGTAGATGATTTGGGGTTGGTCTTGTGGTCACACTTGT 559

528 CCTTGTACATTTCAGAAATAGCCCCCTTCGGAGATTAGAGTTTGCCTGAATACACTACCA 587

560 CCTATTATATCTGAGACTGCACCACCTGAGATTAGGGACTGTTGAATACGCTACCG 619

[illegible][illegible]

046 1CGCCH1CALCCGAI 1GGAGHAI 1A1GCI 1GG1G1GCI 1CGCGA1ACCI 1CA11G11C11C 10

680 ATGCCATCACCTAGCTGGAGATGATGCTGGTGTCTTTTCATCCCTTCCCTTGTCTTT /39

708 TTGGTTGACATATTTTATCTTCCTGAATCTCCAAGATGGCTCGTTAGCAAGGTCGG 767

Db	740	TTGTTCTCAAGGTCCTTCTTCTTGCCAGATCCCAAGGTGGCTGTGACGAAGGTGCA	799
OY	768	ATGGCAGAGGCAGAAAAAGGTGTTGCAAAAATTACGACGGAAAGACATGTTCTCAGGTGA	827
Db	800	ATGCTTTGAAGCAAGCGGGGTTCTTCAGACACACGCGGTGTCGGGAAGATGTGCTGTGTAG	859
OY	828	TTGTCCCTTCCTCTCGAAGAGGTTTGAGAGGTTTGAGAGACATCTTCATTTGAAGGTTCATC	887
Db	860	ATGGCTTTGTGTGTGAGGTCCTTGGAATTGAGAGGTGAACCACTTAAGGAAATATATA	919
OY	888	ATTGGACCTGCCACCGAGCAGCCGATGATCTTGTATCTGACGGTGTATGAAGAACAAATC	947
Db	920	ATTGGTCCCGCGAGTGAATTAATGATGATCATGATATATAGCTGTGATTAAGATCAAAATT	979
OY	948	ACACTTATATGGCCCTGGAAGAAGCCAGTATGATGTCTGCACCTTTCAAGGACCCATC	1007
Db	980	AAGTATATATGGCAGAAAGAGGGCTGAATGTGGTTGTCTAAGGCCAGTCCAAAGA-----	1033
OY	1008	ATGCTTGAAGTGTGCTTTCTCTTGATCATCGTCATGGAGGATGGTGAACCAAGAG---T	1066
Db	1034	-----GGAGACATATAGTGTGTTTGTCTTCGCAATGGAAGTACAAATGACAGAGGCA	1088
OY	1065	GTACCCCTTATGATCCGATTTGACACTTTTGTGATGTCCATGAGAAATATGCTCTCAA	1122
Db	1088	GGCTCATTTGATTGATCTCTCTTGACACATGTTTGGAGGTTTCACGAAAGATGCGCAG	1144
OY	1125	GCTGAGAGAAATATGAGAGACATGTTTCCAAA-TTTTGAAGTATGTTCAAGTGTACAA	1188
Db	1148	ACT-----GAAAGCATGAGAGTGCCTTGTGCCACATTTTGGAGATATGTTCAAGTGTGGA	1200
OY	1185	GATCAGCATGCCAAAATGAGCAGTGGATGGAAGAAATCTTCATAGGGATGACAGAGAG	1244
Db	1205	GGGAATCAACCAAGACATGAATATTTGGATGGAAGAAATCTTGTGTGAAGAAGTGAAGAT	1266
OY	1245	TAGGATCTGATGTGTCAGAGAGGTGAATGAGGACAAATCTCCATAGCCCATTTGCTGTCC	1300
Db	1265	TATCATCCGACCAT---GGAATGATGTTGAAAGATGATCTTCATCTCCGTTGATCTCA	1322
OY	1305	AGGCAGCGAAACAGGTGCGGAGGAAGGAAGCATTTGTSCACATGATGACCGGTGAAGTGT	1366
Db	1322	CGTCAACCGACAAAGCATGGA---GAAAGCATGCTTCACATGCTCATGGAACCTTTCT	1377
OY	1365	TTGACATATGAGAAGCAACCCTCTTAAGGGAGAGGTGAGATAGTGTGACAGACATGAT	1422
Db	1379	ACCTTCAAGACATGGAAGTCAAGTGCAGGGAGCTCAAGGGGAAGGAACGGGTATATGGGG	1433
OY	1425	ATGGGTGGGGATGCGACGTTGCTTGGAAATGCTCAGAGAAGAGTGAATGTGTAGA	1488
Db	1439	ATTGAGGTGATGTGCATGTGCATGGAATATGACGGAAGAGAAATGAATTCGGAGACAG	1499
OY	1485	AAGGAAGGTGTTTCAAAAGAGTCACTCTTGACCAAGAGGAGTTCGGCTCAAGAAAG	1544
Db	1499	AAAGAAAGGTGTTTAAACGAATATCTTGACATCAAGAAAGTTTCCAGAGATCTCAAGT	1555
OY	1545	GGCTCAATTGTTCACTTCCCGGTGTGCGCATGTTTGTGAGGGTATGATGTTGTACAT	1600
Db	1559	GGCTCAATTGTTCAATGCTGTGTGTATGGAACCGGTGAAGCA---GATTTGTACAA	1611
OY	1605	GCTGCTGCTTATGATGATCAGTCAAGCATTTTCTCAAAGGGTCTTGTAACCAAGCATG	1666
Db	1616	GGGTCTGCTTGTGTTAGCCAAACAGCTCTTATTTCCAAAGACCTTCTCAAGAACATACA	1677
OY	1665	TCAGATGCGCATGTTCAACCATGTGAGGTAGCGCAAAAGGTTCAGTTGGAAAGAT	1722
Db	1676	ATTGGTCCCTGATGTGATCATTCATCGAAACACT---AAAGGTCAATTGGCATGAT	1733
OY	1725	TTGTTTGAACCTGAGTGAAGGCGTGCCTGTTAGTGTGTGTGGAATTCAGATCTTCAA	1788
Db	1733	CTTCATGATCTGGAATCAAGGTGATTAAGTGTGAGAGTTGGACTTCAAAATCTTGAG	1799
OY	1785	CAGTTTGTGGAATAAAGGTGTTCTGTACTATATACCTCCAAATTTTGAAGACGCTGGT	1844

Db	1793	CAGTTCACGGCATTCAACGGAGTCTTTACTACACACCGAAATCTTGACGACGGGGT	1852
Qy	1845	GTGGCAGTTATCTTTGCCAAATTTGGTCTACAGCTCGGCATACAGATTCATCTTATCAGT	1904
Db	1853	GTCCGGATCCCTACATATGGAACATGGGGATTAAGTCTTCTCCAGATCTTATATATAGT	1912
Qy	1905	TCTCTCACTACCTTACTAATATGCTCTCTTGACATTTGGCTTTGGCAGTCTATAGATCTT	1964
Db	1913	GCATTGACACCTTTGTGATGTTCCTGCAATACTGTGTGCAATGAGGCTCATGTGATCTT	1972
Qy	1965	TCCGGAAAGAGTTTTTGTGCTAGGACCAATTCGAATCTTGATAGCATCTTAAGTTATC	2024
Db	1973	TCTGGTGGAAAGACCTTGCTTCTCACACGATACCAATCCGATAGCATCTCTATTTGGTT	2032
Qy	2025	CTGTGTGTGTCCATCTTAATTTGATTTGGGTACATAGCCCATAGCTTTGTCTCCACCGTC	2084
Db	2033	TTAGTAATCTCAAACTGTTCACATAGAACGATGTGACCGGGCTTATCAACCGTA	2092
Qy	2085	AGTGTATCGTCTACTCTGCTGCTCGTTATGAGGATTTTGGTCCCATCCCAATTTTAT	2144
Db	2093	AGCGTTGTGCTCTACTTCTGCTTTTGTGTATGGGTTTCGGTCCGTGCCAAACATCTTC	2152
Qy	2145	TGTGCAGAGATCTTTCCAAACGAGGTTCTGTGACCTGTATTTGCAATTTGTGCTTTACA	2204
Db	2153	TGTTCAAGAGATTTTTCCACTCGAGTCCGGGAAATCTGACATCGCATCTGGGCACTACCC	2212
Qy	2205	TTCTCGAGTCGAGATATCATGTGTCACTACACGCTTCTCTGTAGTCTGAATGCTATTTGGA	2264
Db	2213	TTCTCGAGATCTGTGACATATATGTCACTTACAGTCTCCCGTGTGCTCAAAATCCATTTGGA	2272
Qy	2265	CTGGCGGGTGTTTTACAGCATATATGACAGTCTATGCTTGAATTTCTTTGTGTTCGTCCTC	2324
Db	2273	CTAGCTGTGTGTTTGGAAATGTACGCAATGCTATGCTGCAATTTCAATGGGTCTTTGTGTTC	2332
Qy	2325	CTTAAGGTCCCTGAGACAAAGGGAGTGCCTTGAGGTATTAACGAATCTTTTGCAGTT	2384
Db	2333	ATTAAGATCCCGGAAACTAAAGGATGTCACCTTAAGATCAACAGAGTCTTTTCTGTGT	2392
Qy	2385	GGTGCGAAGCAAGCGGCTGC 2404	
Db	2393	GGAGCTACAGCAAGCTGAAAGC 2412	

```
.
RESULT 7
AXS06620          AXS06620      2190 bp       DNA           linear     PAT 27-SEP-2002
LOCUS              Sequence 1315 from Patent WO0216655.
DEFINITION         AXS06620
ACCESSION          AXS06620
VERSION            AXS06620.1   GI:23387857
KEYWORDS
SOURCE
ORGANISM           Arabidopsis thaliana (chale cress)
                   Arabidopsis thaliana
                   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                   rosids; eurosoid II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
AUTHORS            Harper,J.F., Kreps,J., Wang,X. and Zhu,T.
TITLE               Stress-regulated genes of plants, transgenic plants containing same, and methods of use
JOURNAL            Patent: WO 0216655-A 1315 28-FEB-2002;
                   The Scripps Research Institute (US) ; Syngenta Participations AG
                   (CH)
FEATURES
source             Location/Qualifiers
                  .                ..2190
                        /organism="Arabidopsis thaliana"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:3702"
BASE COUNT        529 a    439 c    572 g    650 t
ORIGIN
Query_Match          31.5%; Score 875.8 ; DB 6; Length 2190;
Best Local Similarity 64.9%; Pred. No.2.7e+219;
Matches 1449; Conservative 0; Mismatches 722; Indels 60; Gaps 8;
```

177 ATGTGGGGGCTGTTCTTGTCCGATAGTGGCTCCATGCGCATCTATTTGAGGGGTTG 236
1 ATGAGTGGAGCTGTGCTTGTCTTGTCTATGTGCTGCTGTTGGCACTTTGACAGAGTGG 60
237 GACAAATGCCACCATGCGAGCTGCTGTTCTGTATATTAAGAGAAATTTCAATTTGCAAAAT 296
61 GATAACGCAACTATTTGACAGAGCTGTGTTGATATTAAGAGAGTTTATTTGGAGAGT 120
237 GAGCCCACTGTGAGAGGACTATTTGTCTAATGTCACTTATGCGGCCCACTGTTACT 356
121 AATCATCATGTGAAGGCTTAATTTGGGATGTCACTTAATTTGCTACTCTGTATTTCA 180
357 ACATTTCCCGGGCCATATTCAGACTCGATGCGCCAGCGCTATGCTTATTTCTCTTCA 416
181 ACATCTCTGAGAGGGGTAGCTGATTTGGCTTGGCTCCGCTCCCATGCTAATATTTGCTCA 240
417 ATTCTGTACTTCTTGACGGGCTCATCATGTATGCTCTCTAATGCTATGCTGCTG 476
241 ATTCTCTACTTGTGTTGTTCTCTAGTAATGTATGCTGCTCCGAAATGTTATGTTGCTC 300
477 TTGGAACGCTTGATAGATGATTTGGTATTTGGCTTGGCTGCTCAAGCTTGTGCTTATAC 536
301 TTAGAGAGGTTGTGATGATGATTTGGGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 360
537 ATTTCAGAAATAGCCCTTCCGAGATTTAGAGTTTGTGCTGATATCACTACCAATTTCACT 596
361 ATATCTGAGACTGACACCTGATGATTTAGAGGACCTGTGATATCCGCTACCGACTTCACT 420
597 GGATCAGAGAGATTTCTTGTCTATCATGCTGATGCTGTTGGATGCTCCCTGTGCGCATCA 656
421 GGCTCTGAGAGGATGTTCTTATCTTACTGATGTTGTTGGATGCTGATGATGCTGATCA 480
657 CCCGATTTGAGAAATATGTTGTTGCTGCTGCGATACCTTCAATGTTCTTCTTGGTTTG 716
481 CCTACTGAGATGATGCTGTTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 540
717 ACAATATTTTATCTTCTGATCTCCAGATCTCCAGATGAGTCTGTTAGCAAAAGTGGAGAG 776
541 ACGGCTTCTTCTTGTCCGAGTCCCGAGTGGCTGCTGAGGCAAAAGTGGATGCTTGA 600
777 GCAAAAAGGTTGTCGAAAGTTTACGGGGGAAAGCATGCTCAGTGTGATTTGCTCTT 836
601 GCAAAAGGGGTTCTTCAAGAGCTGGGTGGTGGGAAAGTGTCTGTGATGATGCTTGG 660
837 CTTCTGAAAGGTTGAGGTTGAGAGACACTTCCATTTGAAGATCATCATGCTGACT 896
661 TTGGTTGAGGGTCTTGGAAATGAGGTTGAACCAACCATAGAGAAATATTTATTTGCTCC 720
897 GCCACGAGGAGCGGATGATCTTGTACTGACGGGATGAAGAAATCACTACCTTAT 956
721 GCGGATGAAGTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
957 GGGGCTGAAAGAGCGCATGATGATGCTGCTGACCTTCTTAAGGGAACCCATCATGCTTGA 1016
781 GGTGAGAAAGAGGCTGATGTTGGTTGCTAGGCCAGTCAAGAG-----GGA 828
1017 AGTGTCTTCTTCTGATCTGCTGATGAGAGCATGTTGAACCAAG--TGTACCTCTT 1073
829 AGCATATGATGATGTTGTCTGCGCATGATGATGATGATGATGATGATGATGATGATG 888
1074 ATGATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1133
889 ATTTGATCTCTTGTCACTGTTGAGAGCTTCCAGAGAGATGCGGACCT--GGA 945
1134 AGTATGAGAGCATGTTGTTCCAACTTTGGAAGTATGTTGATGCTCAGATGATGAT 1193
946 ACATGAGAGAGTGTGTTTCCCACTTTTGGAGATGTTTCAAGTTTGAAGAGATCA 1005
1194 GCCAAATATGAGCATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1253
1006 CCAGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1065

1254 GATGTGAGAGAGGATCATATGAGAGCATCTCCATAGCCCATTTGCTCCAGGAGCA 1313
1066 GACCAT---GGAATGATTTCTGAGATGATCTTCAATCTTCCGTTGATCTCAGCTCAACG 1122
1314 ACAGGTGGGAGAGGAGAGCATTTGTGACCAATGTGACCGTGAAGTCTTTGAGCATG 1373
1123 ACAGCATGGA---GAAAGCATGCTCAGCTGCTCATGGAACCTTTCTACCTTCAGA 1179
1374 AGAAGCAAAACCTTTTGGGAGGATGATGATGATGATGATGATGATGATGATGATGATG 1433
1180 CATGAAATCAATGATGAGAGGATCTCAAGGAGAGGAGGATGATGATGATGATGATGATG 1239
1434 GGATGCACTTGTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1493
1240 GATGCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1299
1494 GGTTCGAAAGATCTACTTSCAACAGAGGAGATCTCTGCTCAAGAGGGGCTCAATT 1553
1300 GGTTC-----CCAGATCTGACGTTGCTCAATT 1329
1554 GTTTCACTTCCCGGTGGGAGATGTTTGAAGGATGATGATGATGATGATGATGATGATGAT 1613
1330 GTTTCACTTCCCGGTGGGAGATGTTTGAAGGATGATGATGATGATGATGATGATGATGAT 1386
1614 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1673
1387 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1446
1674 GGCATGTTCAACCATCTGAGTATGCTGCAAGAGGATGATGATGATGATGATGATGATGAT 1733
1447 GCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1503
1734 CTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1793
1504 CTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1563
1794 GGAATGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1853
1564 GGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1623
1854 ATTTCTTCAAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1913
1624 CTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1683
1914 ACTTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1973
1684 ACTTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1743
1974 AGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2033
1744 AGGACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1803
2034 TCCATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2093
1804 TCAATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1863
2094 GTTACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2153
1864 CTTACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1923
2154 ATCTTCAACAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2213
1924 ATTTTCAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1983
2214 GAGATATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2273
1984 TGTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2043
2274 GTTTTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2333
2044 GTTTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2103
2334 CTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2393

RESULT 8
ATH532570 2190 bp mRNA linear PLN 11-DEC-2002
LOCUS Arabidopsis thaliana mRNA for monosaccharide sensing protein 2
DEFINITION (msp2 gene).
ACCESSION AJ532570.1 GI:26800694
VERSION monosaccharide sensing protein 2; msp2 gene.
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustroids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Stamme, C., Tjaden, J., Trentmann, O., Emmerlich, V. and Neuhaus, E.
TITLE A novel family of plant monosaccharide transporters is involved in a new type of eukaryotic sugar sensing
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2190)
AUTHORS Neuhaus, E.
TITLE Direct Submission
JOURNAL Submitted (04-DEC-2002) Neuhaus E., Plant Physiology, University of Kaiserslautern, Erwin-Schroedinger-Str. 22, 67663 Kaiserslautern, GERMANY
FEATURES
source Location/Qualifiers
 1..2190
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 1..2190
 /gene="msp2"
 1..2190
 /gene="msp2"
 /codon_start=1
 /product="monosaccharide sensing protein 2"
 /protein_id="CAD58692.1"
 /db_xref="GI:26800694"
 /translation="MSGAVLVAAVGLLQGMNATVAGAVLYIKKEFNLESNPV
 EGLVAMSLIGATLITTCGSAVADWLGRIPLLSILYFVSIVLMSPNVYVLLG
 RLIDRGVGLVTVLPIYISERAPPEIGLNTLPQFGSGMPLSYCGRMSLMPS
 PSMRLGVLPVLPVPEPLVPELPESPRMLVSKGMLEAKRVLORLGRBVDVGEN
 ALLVGLIGETTTTEETIIEYIPADRYTDHDAVDKQIKLYGAEGLSWARPVKGG
 STMSVLSRHSSTMSRRQSLDPLVTLRGSVHKMPDTGSMKSALEPHGSPFVGN
 QPRHMDWBEINLVGEDEYPSDHGDSDSDLSPLISQTSMEQMDHPATGLSTF
 RHGSQVQAGAGSMGIGGQVAMKTEREDSEKBEFPGSRGSIYSLPGD
 GTGEADPVOASALVSPALYSKDLKEHIGPAMVHPSETTGSIMHLDHPGVRAL
 VVVGLOILOPSGNGVLYYRPOILDOAGVGIILSNMGISSSASLILSTTFVNL
 PAIYAMKMLDSGRRLLITPIPLIASLVLVYLSNVHMSIYHVALISTSVLYE
 CFFVWGPAPNITLSELPTRKRGICAIACALTCALTCIDIIYVSLPALLSLGIAGV
 FGMVAIVCCISNVFVIRKPEKMPLEVTTEFFSVGARQDAANNE"
BASE COUNT 529 a 439 c 572 g 650 t
ORIGIN
 Query Match 31.5%; Score 875.8; DB 8; Length 2190;
 Best Local Similarity 64.9%; Pred. No. 2.7e-219;
 Matches 1449; Conservative 0; Mismatches 722; Indels 60; Gaps 8;

297 GAGCCACTGAGGAGCACTAATTGTGCATATGCTACTTATCGGCGCCACCATGTTACT 356
 121 AATCATATGAGAGGCTAATTTGTGCGAGATGACATTAATGATGCTATGATTAACA 180
 357 ACATTTCTCGGGCCATTATCAGACTGATTTGGCCGAGCCCTATGCTTATCTCTTCA 416
 181 ACATGCTCTGGAGGGGATGATGATTTGGCTTGTGGCCGTCCATGCTAATATTCCTCA 240
 417 ATCTGTACTTCTTACGCGGCTCATCATGCTATGCTCTCTAATGCTATGCTCTGTG 476
 241 ATCTCTACTTGTGTGTTCTCTGTATGCTATGCTCTCGAATGTTATGTTGTGCTC 300
 477 TTGGACGCTTCGTATGATGATTTGGTATTTGGCTTGTGCTGACGCTTGGCTTGTAC 536
 301 TTAGAGGTTGTATGATGATTTGGGGTGTCTTGGTACACATGTTGCTATTTAT 360
 537 ATTTCAGAAATAGCCCTTCGAGATTTAGAGTTTGTGTAATACATACCAATTCAGT 596
 361 ATATCTGAGACTGACACCACTGATATTTAGGGAGCTGTTGATATGCTACCGCATCT 420
 597 GATTCAGAGAAATGTTCTGTCTCATCTGATGCTATGCTGTTGGATGCTCTGCTGCA 656
 421 GGCTCTGAGAGGATGTTCTTACTGTATGATGTTTGGAAATGCTGTATGATCA 480
 657 CCCGATTGAGAAATTAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 716
 481 CTTAGCTGAGATGATGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 717 ACAATATTTTATCTTCTGATCTCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 776
 541 AOGGCTCTTCTTGGCCGAGATCCCAAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTG 600
 777 GCAAAAAAGTGTGCAAAATTAAGGGGAAAGACATGCTCTGAGTGAATTTGCTT 836
 601 GCAAGCGGGTCTTCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 837 CTTCTCGAAGGTTGAGGCTTGAAGAGACATTCCTCATGTAAGTATCATGATGGAAT 896
 661 TTGGTGGGCTTGTGATTTGAGTGAAGTAAACAACATGAGGAATATATATTTGCTCC 720
 897 GCCACCGAGCGCGATGATCTTGTACTGACGCTGATTAAGAAACAATCACTTAT 956
 721 GCGATGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 957 GGGCTGAAGAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1016
 781 GGTGAGAAAGAGGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 828
 1037 AGTGTCTTCTTCTGATCTGCTGATGAGGAGATGAGAAACAAG---TGTACCCCT 1073
 829 AGCATATGAGTGTGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 888
 1074 ATGATCCGATTTGATCACTTTTGTAGTGTGCTGATGAGAAATATGCTCAAGCTGAGA 1133
 889 ATGATCTCTTGTACACTGTTTGGAGCGTTACAGAAAGATGCGGACACT---GGA 945
 1134 AGTATGAGACCAATTTTCCAACTTTGGAAGTATGTTCACTGATGATGATGAT 1193
 946 AGCATGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1005
 1194 GCCAAAAATGAGAGTGGATGAGAGAAATCTTCAATAGAGTGAAGAGATGATGAT 1253
 1006 CCAAGACATGAAAGTTGGATGAGAGAAATCTTGTGAGAAAGTGAAGATTAATCAATC 1065
 1254 GATGCTGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1313
 1066 GACCAT---GGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1122
 1314 ACAGGTGGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1373
 1123 ACAAGCATGGA---GAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1179
 1374 AGAAGCAAAACCTCTTATGAGGAGGATGAGATGATGATGATGATGATGATGATGAT 1433

Db 1180 CATGAGACTCAAGTCCAGGAGGCTCAAGGGGAGAGGAGGGAGTGGAGTGGAGT 1239
 Qy 1434 GGATGGAGCGCTTGTGAAAAATGTCAGAGAGAGGAGAGATGATGAGAGAGAGT 1493
 Db 1240 GGATGGAGCGCTTGTGAAAAATGTCAGAGAGAGGAGAGATGATGAGAGAGAGT 1299
 Qy 1494 GGATGGAGCGCTTGTGAAAAATGTCAGAGAGAGGAGAGATGATGAGAGAGAGT 1553
 Db 1300 GGTTC-----CCAGATCTCGAGCTGGCTCAATT 1329
 Qy 1554 GTTTCACCTCCCGGTGGTGGCGATGTTTGAAGGAGTGAAGTGAAGTGAAGT 1613
 Db 1330 GTTTCACCTCCCGGTGGTGGCGATGTTTGAAGGAGTGAAGTGAAGTGAAGT 1386
 Qy 1614 TTAGTAAGTCAAGTCAAGTCTTTCAGAGAGTCTTTCAGAGAGTCAAGTCAAGT 1673
 Db 1387 TTGTTAGTCAAGTCAAGTCTTTCAGAGAGTCTTTCAGAGAGTCAAGTCAAGT 1446
 Qy 1674 GCGATGTTCAAGTCAAGTCTTTCAGAGAGTCTTTCAGAGAGTCAAGTCAAGT 1733
 Db 1447 GCGATGTTCAAGTCAAGTCTTTCAGAGAGTCTTTCAGAGAGTCAAGTCAAGT 1503
 Qy 1734 CCGTGAAGTCAAGTCTTTCAGAGAGTCTTTCAGAGAGTCAAGTCAAGTCAAGT 1793
 Db 1504 CCGTGAAGTCAAGTCTTTCAGAGAGTCTTTCAGAGAGTCAAGTCAAGTCAAGT 1563
 Qy 1794 GGAATTAAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAG 1853
 Db 1564 GCGATGTTCAAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAG 1623
 Qy 1854 ATCTTTCAGAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAG 1913
 Db 1634 CTACTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1683
 Qy 1914 ACCTTACTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 1973
 Db 1684 ACCTTACTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 1743
 Qy 1974 AGTTTGGTCTGAGAGAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAGAGAG 2033
 Db 1744 AGGACCTGCTTCTCAAGAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAGAG 1803
 Qy 2034 TCCATCTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 2093
 Db 1804 TCCATCTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 1863
 Qy 2094 GTTACTGCTGAGAGAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAGAGAGTCT 2153
 Db 1864 CTCTACTGCTGAGAGAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAGAGAGT 1923
 Qy 2154 ATCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAG 2213
 Db 1924 ATCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAG 1983
 Qy 2214 GGAATTAAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAGAG 2273
 Db 1984 GGAATTAAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAGAG 2043
 Qy 2274 GTTTCAGAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAGAG 2333
 Db 2044 GTTTCAGAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAGAG 2103
 Qy 2334 CCGTGAAGTCAAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAG 2393
 Db 2104 CCGTGAAGTCAAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAG 2163
 Qy 2394 CAGAGAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAGAG 2404
 Db 2164 CAGAGAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAGAGAGTCTTTCAGAG 2174

RESULT 9

HVU534446
 LOCUS HVU534446 2516 bp mRNA linear PLN 14-FEB-2003
 DEFINITION Hordeum vulgare mRNA for sugar transporter (sugp2 gene).
 ACCESSION AJ534446
 VERSION AJ534446.1 GI:26986187
 KEYWORDS sugp2 gene; sugar transporter.
 SOURCE Hordeum vulgare subsp. vulgare
 ORGANISM Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Pooideae; Triticaceae; Hordeum.
 REFERENCE
 1 Weschke, W., Panitz, R., Gubatz, S., Wang, Q., Radchuk, R., Weber, H. and Wobus, U.
 The role of invertases and hexose transporters in controlling sugar ratios in maternal and filial tissues of barley caryopses during early development
 Plant J. 33 (2), 395-411 (2003)
 JOURNAL MEDLINE 22424051
 PUBMED 12535352
 REFERENCE 2 (bases 1 to 2516)
 AUTHORS Radchuk, R.
 TITLE Direct Submission
 JOURNAL Submitted (09-DEC-2002) Radchuk R., Molecular Genetics - Gene Expression, Plant Genetics and Crop Plant Research, Corrensstr. 3, Gatersleben, D-06466, GERMANY
 FEATURES
 source location/Qualifiers
 1..2516
 /organism="Hordeum vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cultivar="barke"
 /db_xref="taxon:112509"
 /tissue_type="developing caryopsis"
 /clone_lib="lambda Zap2"
 /dev_stage="0-15 days after flowering"
 1..2516
 /gene="sugp2"
 77..2338
 /gene="sugp2"
 /codon_start=1
 /product="sugar transporter"
 /protein_id="CAD58959.1"
 /db_xref="GI:26986188"
 /translation="MRGGRSMRGGRGMRARSRRLRGRRRRAGAGARGLAA
 LQAKMSIIEBEIYMAAAKLVFLASVYMLAPVYLLARLDGICGLATVPL
 YISFAPDILKGLNTLPFGSGGMLSTCMTVPLMPPDRIMGVLSISLMT
 FALTVFLPESPRMLVSKGRMAEKRVLQRLGRDVSGEMLALVEGLGVKTHIEY
 IIGDDDELADGLAPDQKLYGAEBSVTLAPRVGGQSALGSLGMSRHSW
 SGRSLVDPLVTLFGSVHEKMEVMSNRSTLPNFGSMFSAEQOQAKMDAESH
 DDERYASDHGADDIEDNINSPLISROATSVGEKELIAPHGIMGVSSSMVGDDAYS
 SMGICGQMLAKRTTERBGRDGRGGAQRILYHEBSVSDRRSISLMPGDI PRG
 EYIDAAALVSGPALYSKDLLEQULAGPAMVPSAVKGVKMAELFEPGVHAFVLG
 GLQILQOFAGINGLVYTPQILEQAGVGIILSNGLSSSASILISALTLLMGLYS
 DRIGARLTAASPMYCSIPALFYPRAILVIVNLVDGTVMHALSTISVIVCFEV
 MGPIPEGSTSVKIDTDRAANAASLAKLLVFWICDILVTVLTPVNLNIGLAGEGYI
 AVCMKIAFVPMYVMPETKMPLEVIIEFEPVGAKKQKEATDEVAEKIAESRPEBEK
 SPVDYVDHVMKV"
 BASE COUNT 567 a 524 c 774 g 651 t
 ORIGIN
 Query Match 29.0%; Score 806.2; DB 8; Length 2516;
 Beet Local Similarity 65.1%; Pred. No. 6.1e-201;
 Matches 1294; Conservative 0; Mismatches 668; Indels 27; Gaps 6;
 Qy 415 CAATTCTGACTTCTTCAGCGGCTCATGCTAGTCTCTTAATGCTATGCTCTGC 474
 Db 264 CGAACTCGTCTTCTTCGTCGCAACGATGCTCTGGGCGCCCAACGCTATGCTGC 323
 Qy 475 TGTGGACGCTTGTGATGATGATTTGGTATGGCTTGGCTGTACCGTTGGCTTGT 534
 Db 324 TCTGGCCAGGCTATCGACGGGTTCGTTAGCTGGTGGAGTCAACCTTGTCTCTGT 383
 Qy 535 ACATTGAGAAATAGCCCTTCGAGATTTAGAGTTTGCTGAATACACTACCAATTC 594

Db 384 ACATTTCGAGACCCGCCGACCTGACATTAGAGGGCTGCTGAACACGCTCCGCGACCTTCA 443
 Qy 595 GTGATCAGAGAGAAATGTTCTTGATACATGATGATGTTGGAGATGTCCTGCGCAT 654
 Db 444 GTGGGTCAAGAGAGATGTTCTTTCTTACATGATGTTTACATGTCCTCCTCAGCGCC 503
 Qy 655 CACCCGATTTGAGAAATATGCTTGTGTGCTGCGCATACCTTCACTTGTCTTCTTGGTT 714
 Db 504 AGCCGACCTGAGAAATCATGCTGGGGGTTTGTGATCCCGTCCGCTTAATGATTTTTCAT 563
 Qy 715 TGAATATATTTATCTTCTGATCTCAAGATGCTGCTTGAAGAAAGGTCGGATGCGAG 774
 Db 564 TGACGTCTTTTATTTGCCCCGAACTCCAGATGCTGTGTGAGCAAGGAAAGATGCTG 623
 Qy 775 AGGCAAAAAAGTGTTCGAAAAGTTACGGGGGAAAACAGATGCTTCAGTGAATTTGCC 834
 Db 624 AGGCCAAGAGATATTTGCAAGACGTCGGGGAAGGAGATGCTCAGAGAGATGCGCC 683
 Qy 835 TTCTTCTGAGAGGTTTGAAGGTTGAGAGACATCTTCCATTTGAAGATCATATTGAGC 894
 Db 684 TTCTTGTGAGAGATTTGGGTGTT---GGAAAGACACATTTGAGGATTCATTAATTTGAC 740
 Qy 895 CTGCCACGAGGAGCCGATGATCTTGTACTGACGATGATGAAGAACAAATCACACTT 954
 Db 741 CTGATGATGAGCTGCTGATGACGGGCTGCTCCA---GATCAAGAGAGATTGAACCTT 797
 Qy 955 ATGGGCTGAGAGAGGCCAGATGATGATGCTGACCTTCTAAG-----GGACCATCA 1008
 Db 798 ATGAGCTGAGAGAGGGGATCTTGGATGCGCCGTCCTGTAGGGGTGCGGCCAAGTG 857
 Qy 1009 TGCTTGAAGTGTCTTCTTCTGATCTGCTCATATGAGAGCATGCTGAACCAAGATGAC 1068
 Db 858 CACTTGAAGCCCTTTGGGCTCATGCTGCTCATAGAGAGATGCTTCTGAGGGTAAAT 917
 Qy 1069 CCCTTATGATCCGATTTGATGACATCTTTTGTAGTATCATGAGAAATATGCTCAAGCTG 1128
 Db 918 CTCTTGTGACCCCTTGTGACCTCTTCTTGAAGATTCATAGAGAGATGCTGAGGTA 977
 Qy 1129 GAGAGATATGAGAGACATTTTCCAACTTTGAGAGATATGTTCACTGTCAGAGATC 1188
 Db 978 TGCGGAGATGCGGAGACCTTGTCTTAACCTTGGAGAGATGTTAGTGGCTGAAC 1037
 Qy 1189 AGCATGCCAAATATGAGAGTGGATGAGAGAAATTTTATAGGATGACGAGAGTACG 1248
 Db 1038 AGGACAGGCTTAACCTGACTGGAGATGCTGAGA---GTCATAGGAGATGATGAATTTAG 1094
 Qy 1249 CATCTGATGTCAGAGAGGTCATAGAGACATCTCCATGACCCTGCTGTCAGGCT 1308
 Db 1095 CATCGATCATGTGCTGATGATGATGAGATTAACCTCAATAGCCCGCTTATTTCTGCTC 1154
 Qy 1309 AGGCAACAGTGCAGAGAGGAGAGACATTTGTGACCATGATGACCGTGAAGTGTCTTGA 1368
 Db 1155 AAGGCAAGAGCTGAGGGGTAAGAGATTTGTGACCTCATGGAAGCATTAATGGGT--- 1210
 Qy 1369 GCATAGAGAGCAACCTCTTAAGGGAGGGGTGAGATGCTGTGACAGACATGATATG 1428
 Db 1211 ---GGGTGAGAAATGATGATGATGACAGGAGGGGACCAATAGACACATGAGCATTTG 1265
 Qy 1429 GTGGGAGATGAGAGCTGCTTGAATGCTCAGAGAGAGAGAGTGAAGTGGTGAAGAG 1488
 Db 1266 GTGGGGGTTGGAGATTTAGCTTGAAGTGAATGAGAGAGAGAGTGAAGTGGGCAAAAG 1325
 Qy 1489 AAGGTGTTTCAAAAGATCTAATGTCACCAAGAGAGAGTTCCTGCTCAAGAGGGGCT 1548
 Db 1326 AAGGTGGGCGCCAGGATTTAATCTTGCATGAGAGAGGGGTGTCAAGTATGGAAGGGCT 1385
 Qy 1549 CAATTTTCACTTCCCGGTGTGCGAGATTTTGAAGGTAGTGAATTTGATCATGCTG 1608
 Db 1386 CTATTTTGTCTATGCAAGAGAT---GATATTTCTCTGCTGTGTGATATATCCAGGCG 1442
 Qy 1609 CTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1668

Db 1443 CCGCTTGTGAGCAACCTGCTCTTACTGAGAGACCTGATAGAGCAGGCTTGTG 1502
 Qy 1669 ATGCTCCATGATTCACCATCTGAGTACTGCCAAAGTTTCACTTGAAGATTTGT 1728
 Db 1503 GTCCAGCATGATCATCATCCAGAGGATGCGAAGGGGTAAAGAGGGGCGAGAACTAT 1562
 Qy 1729 TTGAACCTGAGAGAGGCGGCTGCTGATGCTGAGTGTGAGATTCAGATCTTCAAGCT 1788
 Db 1563 TTGAACCTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1622
 Qy 1789 TTGCTGAATTAACGCTGCTGCTGATTAATTAACCAAAATTTGAGCAAGCTGCTG 1848
 Db 1623 TTGGGGGATTAACGAGATGCTGCTGATTAATTAACGAGCTTGAATCTGAGCAAGGCTG 1682
 Qy 1849 CAGTATTTCTTCCAAATTTGCTGAGCTGCGATCAGCATCTTATCACTTCTC 1908
 Db 1683 GGATTTCTTATCAAACTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1742
 Qy 1909 TCACTACTTACTAATGCTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1968
 Db 1743 TGACAACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1802
 Qy 1969 GAAGAGGTTTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2028
 Db 1803 CTGAGAGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1862
 Qy 2029 TTGTGCTAATCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2088
 Db 1863 TTTGTGTAATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1922
 Qy 2089 TTATGCTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2148
 Db 1923 TCATGATCTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1982
 Qy 2149 CAGAGATCTTTCACCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2208
 Db 1983 TCAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2042
 Qy 2209 GGATGCGATATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2268
 Db 2043 GGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2102
 Qy 2269 CGGGTGTTCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2328
 Db 2103 CTGAGATTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2162
 Qy 2329 AGTCTCCTGAGCAAAAGGAGTCCCTTGAAGTATTAACGAATCTTTCAGTGTG 2388
 Db 2163 AGTCTCCTGAGCAAAAGGAGTCCCTTGAAGTATTAACGAATCTTTCAGTGTG 2222
 Qy 2389 CGAAGCAAG 2397
 Db 2223 CGAAGCAAG 2231

RESULT 10
 AX412656
 LOCUS AX412656 2205 bp DNA linear PAT 14-JUN-2002
 DEFINITION Sequence 420 from Patent WO0222675.
 ACCESSION AX412656
 VERSION AX412656.1 GI:21445114
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Burkavota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE
 1 Glazebrook, J., Wang, X., Dangl, J.L., Bulgem, T. and Zhu, T.
 Plant genes, the expression of which are altered by pathogen
 infection
 Patent: WO 0222675-A 420 21-MAR-2002;
 Syngenta Participations AG (CH); UNIVERSITY OF NORTH CAROLINA AT

CHAPL HILL (US) ; Glazebrook, Jan (US) ; Wang, Xun (US) ; Dangl,
Jeffrey L. (us) ; Bulgem, Thomas (US)
Location/Qualifiers
1..2205
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"

BASE COUNT 521 a 524 c 521 g 639 t
ORIGIN

Query Match 25.2%; Score 700; DB 6; Length 2205;
Best Local Similarity 61.2%; Pred. No. 66-173;
Matches 1371; Conservative 0; Mismatches 770; Indels 99; Gaps 11;

177 ATGTGGGGGCTGTTCTTTCGCCAATGCTGCTCCATGCGAATCTATGGAAGGGTGG 236
1 ATGGAAGGAGCACTCTGTTGCTCGCGCCACATGCGCAATTTCTTCAAGAGATG 60
237 GACAAATGCCACATCGACAGCTGTTCTGTATATATAAGAAATTTCAATTGCAAAAT 296
61 GACATGCCACATCTGAGCTATGTTATATACAAAGACTTGAACTTACCA-- 118
297 GAGCCACTGTGAGGAGACTAATTTGTCAATGTCACTTATCGGCGCACATCGTTACT 356
119 ----CCTGTTCAGAGTCTTGTCGTGCTATGTCAATGATGCGTGCAACGGTCAACG 174
357 ACATTCCTGGGCGCATATACACTGATTTGGCCGAGCGCCATAGTTTATCTCTTCA 416
175 ACTGTGTCAGAGCCGATATCTATGCTGCGAGAGCGCCCAATGCTCAATTTATCATCA 234
417 ATTCTGTACTCTTTCAGCGGCTCATCATGCTATGCTCTCTTAATGCTATGCTGCTG 476
235 GTTATATATTCGTCGCGGTTGATATGTTGTGTGCTCCCAATGTCTATGTTCTGTC 294
477 TTGGCAGCGTTCGTAGATGATTTGGTATTTGGCTTGGCTGTCACGCTTGTCTTTGAC 536
295 TTGTGAGGCTTCTTATGAGGTGTGCTGCGGCGCTGCGGTACACTTGTCTCTGTTTAC 354
537 ATTTCGAATATAGCCCTTCGAGATTTAGAGTTTCTGATATACATTAACAAATTCAGT 596
355 ATTTTCGAACCGCTCTCCGAGATCAAGAGACATTAATACTCTCCCTCAAGTTTCTT 414
597 GGATCAGAGAGATTTCTTGTCTACTGCTCATGCTGTTTGGAGTCCCTGTGCGCATCA 656
415 GCGTCTGCTGAGATGTTTGTGTCTATGATGTTTCTATGATGCTCTGAGTGTCTC 474
657 CCCGATTTGAGATTAATGCTGTGTGCTGCGCATCACTTCAATGTTCTTCTTGTGTTG 716
475 CTTAGCTGAGAGCCATGCTCGGTGCTCTCGATCCCTTCTTATTTGTTCTC 534
717 ACAATATTTTATCTTCTGTAATCTCAAGATGCTGTTAGCAAGAGTGGATGCGAAG 776
535 ACGGTGTTTATTTTTCGCGAGTCTCTGTTGCTGTTAGTAAAGAAAGATGACGAG 594
777 GCAAAAAGGTTGCAAAAGTTACGAGGGAAGAGATGCTCAGATGATGTTCCCTT 836
595 GCTACAGAGTTTCTTAACAGTTATGTGCGAGAGAGATTTACCATGATGAGCTTTA 654
837 CTCTCTGAGAGGTTGAGGTTGAGAGACATTTCAATTGAGAGATCATTTGACCT 896
655 CTAGTTGAGAGACTAGATATAGAGAGAAAAAACAATGAGATCTCTTAGTACTTTG 714
897 GCCACGAGGACCGCATAT--CTTGTATCTGACGCTGATAGAAACAAATCAACTT 953
715 GAGGATCATGAAAGTATGATACATTTGAAACGTTGATGAGATGACAAAGCGGCTT 774
954 TATGGGCTGGAAGAGCGCATGATGATGCTGCACTTCTTAAGGAGCCCATCATGCTT 1013
775 TATGGAACCAAGAGATCATGATCTTGTGCTAGACCTGTCCCAAGACAA----- 826
1014 GGAAGTGTGCTTCTCTTGATCTCTGATGAGAGATGAGTGAACCAAGTGAACCTT 1073
827 -ATAGCTCACTTGGGCTACGCTCTGCGCAGGAAGCTTAGCAAAACAAAGCATGATCTT 885

1074 ATGATCCGATTTGTGACACTTTTGTAGTGTCCATGAGATATATGCTCAAGTGAAGA 1133
886 AAAGATCCGCTCTGTCATCTTTTGGCAGTCTCCACAGAGAAAGATCCAGAAAGCGCGGA 945
1134 AGTATGAGAGACCATGTTTCCAACTTTGGAAGTATGTTTCAG-----TGTACAGAT 1187
946 AAGACTCGAGTGGGATTTTCCCTCATTTCCGAGACATGTTCAATGATCTAGCCGATGCG 1005
1188 CAGATGCCAAAAATGACATGGGATGAGAG-----ATCTTCAATAGGATGAC 1238
1006 CCTACCGGTAAACCGGCTCATTTGGAAAGAGACATGAGACCATTCACAAAGACAT 1065
1239 GAGGATGCGATCGATGATGATGAGAGAGGATGATGAG-----ACATCTCCAT 1289
1066 GATGACTATGCGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1125
1290 AGCCATTTGCTGTCCAGAGCAGCAACAGTGCAGAGGGAAGGAAACATTTGACCATGCT 1349
1126 AGCCCTTATATGTGCGCGCAGACCAAGCATGGA---CAAGATATGATCCCATCT 1182
1350 CACGCTGGAAGTCTTTGAGCATGAGAGGAAACCTCTTTAGGGAGGCTGAGATGCT 1409
1183 ACAAGTGAACACATTTAAGCATGAGACACAGTACGCTTATGCA---AGGACACGC 1239
1410 GTGACACACATGATATGCTGAGGAGATGAGCAGCTTGCTTGAATGCTCAGAGAGAA 1469
1240 GAAGTACACATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1279
1470 GGTGGAATGATGAGAAAGAGTGTCTTCAAAAGATCTATGACCAAGAGGAGT 1529
1280 -----ATAGTACGAAGAAAGATGATATCAAGAGATATATCTTAAAGAGATGA 1329
1350 CTTGCTCAAGAAAGGCTCAATTTGTTCACTTCCCGGTGCTGAGATGTTTGAAGGT 1589
1330 GCTGATCTCGCGGCTGATCATATCTTATCCGAGGATCGGAGTGT-----GGA 1383
1590 AGTATGTTGATCAATGCTGCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1649
1384 GCGACTACATTTACAGCTTCTGCTTGAAGAGATGATGATGATGATGATGATGATGATGATGAT 1443
1650 GCTGAACCAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1709
1444 CATGA-----TCGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1488
1710 TCACTTGAAGAAATTTGTTGAACCTGAGTGAAGGCTGCTGTTAGTGTGTTGA 1769
1489 CCACTCTGCTGCTCTTCTTGAACCTGCTGTTAAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1548
1770 ATTGATCTTCAACAGTTTGTGAGTAAACGCTGTTCTGATATATACCAAAAT 1829
1549 ATTCAAAATGCGACAGATTTTCAAGTATCAATGAGTCTTCTTACTTACATCTCAGATT 1608
1830 CTGGAAGAGCTGTGTGAGCAATTAATCTTCCAAATTTGATCTTCAAGCTCGGATCAGCA 1889
1609 CTCGAACGGGCTGGGTGATATTTCTTTCGAGCCTCGAGCTAAGTTTCAATCTCTGCG 1668
1890 TCCATCTTGAATGATTTCTCTCACTTACTTATGCTTCTTCTGATGATGCTTGTGCTGCTGCT 1949
1669 TCATTCCTCATCAGCGGTTTAAACAATTAATCAATGATGATGATGATGATGATGATGATGATGAT 1728
1950 CTGCTTATGATCTTTCGGAAGAGGTTTGTGCTGATGAGCAAAATTCATCTTGAFA 2009
1729 AGACTCATGATGATATTCGGAAGAGGATGATTTACTTCTTGAACATCTCAGATTCTTCAAT 1788
2010 GCATCTTATGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2069
1789 GTCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1848
2070 TTGCTCTCAGCGTGAAGTATGCTTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2129
1849 GCATCTTCAACAGGTTGTGTGCTGCTTCAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1908

QY 2130 ATCCCAACATTTTATGTCAGAGATCTTCCACGAGGTTGTCGCTCTGTATGTC 2189
DB 1909 ATTCACAAACATCTCTGTTCTAAATCTTCCAAACAGATCCGTCGTCTGCAATCGCC 1968
QY 2190 ATTTGTCCTTACATCTGATATCGAGATATCATGTCACCTTACAGCTTCCTGTATG 2249
DB 1969 AATATGTCATGTCCTTTTGGATGTCAGACATTTATGTACGATCTCACTTCCGTTCTC 2028
QY 2250 CTGAATGCTATTTGACGTCGCGGCTTTTTCAGCATATATGACAGTGTATGCTTATTC 2309
DB 2029 CTCAGCTTCATGTCAGTATGTCGTTTTCAGCATTTACGTCGCGTTTGCCTTATCTCA 2088
QY 2310 TTTGTCGTCCTCTTAAAGTCCCTGAGACAAAGGATGCCCTTGAAGTTATTC 2369
DB 2089 TGGATCTTGTTCATGAAAGTCCCGAGACCTAAAGCATGCTTTGAGATATACCA 2148
QY 2370 GAATCTTTGACAGTGTGTC 2389
DB 2149 GACTACTTTGCTTGGAGC 2168

RESULT 11
AX507559 2205 bp DNA linear PAT 27-SEP-2002
LOCUS AX507559 Sequence 2254 from Patent WO0216655.
ACCESSION AX507559
VERSION AX507559.1 GI:23388796
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Bukariyola: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots;
rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
AUTHORS Harper, J.F., Kreps, J., Wang, X. and Zhu, T.
TITLE Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
JOURNAL Patent: WO 0216655-A 2254 28-FEB-2002.
The Scripps Research Institute (US); Syngenta Participations AG
(CH)

FEATURES
source 1..2205
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"

BASE COUNT 521 a 524 c 521 g 639 t
ORIGIN

Query Match 25.2%; Score 700; DB 5; Length 2205;
Best Local Similarity 61.2%; Pred. No. 66-173;
Matches 1371; Conservative 0; Mismatches 770; Indels 99; Gaps 11;

QY 177 ATGTGGGGGCTGTTTGTGTCAGATGTCGCTTCATGCAATCTATTTGAGGGGTGG 236
DB 1 ANGAAGGAGGACACTCTGTCGTCTGCGCCGACATGCGCAATTTCTTCAAGAGATG 60
QY 237 GAGAATGCCACCATGTCAGCTGCTGTTCTGTATATTAAGAGAAATTTCAATTCGAAAT 296
DB 61 GACATGACACCATGTCAGCTATGTTTATATCAACAAGACTGTAATCTACCAA- 118
QY 297 GAGCCACGTGTGAGGAGTAATTTGTCATATGTCATTTAGGGCCCATCGTTACT 356
DB 119 ---CCTCTGTTCAAGGTCCTGTCGTATGTCATGATGTCGCAACGGTCATCAAG 174
QY 357 ACATTCCTCGGAGCATTTATCACTGATGTCGCGACGCGCTATGCTTATCTCTTCA 416
DB 175 ACTTGTGACGACCATATCTGATTTGCTGCGACAGCGCCCATGCTCATTTATCATCA 234
QY 417 ATTTGTAATTTCTTACGCGGCTCATGATGATGATGTCCTCAATGTCATGCTGCTG 476
DB 235 GTATGTAATTTCTGTCGCGTTGATGATGTTGTGATCTCCCAATGTCATGTCGTGTC 294
QY 477 TTGGACGCTTGTGATGATGATTTGATTTGGCTTGTGCTGTCACGCTTGTGCTTGTAC 536

DB 295 TTTCGTAGGCTTCTTATGATGATTTGTCGCGGCTCCGCGTTACATCTTGTCTTAC 354
QY 537 ATTTGAGAAATAGCCCTTCGAGATTTAGAGTTTGTGTAATACATACCAATTCAGT 596
DB 355 ATTTCTGAAACCGCTCTCCGAGATGAGAGACAGTTAAATATCTCTCCAGTTCTT 414
QY 597 GATTCAGAGAAATGTTCTTGTATCTGATGATGATGTTTGGAGATGTCCTGTCGATCA 656
DB 415 GGCCTGTGTGAATGTTTGTCTATGATGATGATGTTTCTATGTCCTGATGATGATC 474
QY 657 CCCGATGAGAAATATGCTGTCGTCGAGATCTTCAATGTTCTTCTTGTGTTG 716
DB 475 CTGATGAGAGACATCTGCTGCTGTCGATGCTTCTCTTATTTGTTCTC 534
QY 717 ACAATATTTATCTTCTGTAATCTCAAGATGCTGTTAGCAAAAGTGGATGAG 776
DB 535 ACGGTGTTTATTTGCGGAGTCTCTGCTGTCGCTGTTAGTAAAGAAATGAGAGAG 594
QY 777 GCAAAAAGTGTTCGCAAAATTTACGGGGGAAAGCATGTCCTGAGTGAATTTGCTT 836
DB 595 GCTAAGCAGATTCCTTCAACAGTTATGTCGAGAGAAAGTGTACCATGATGATGCTT 654
QY 837 CTTCGCAAGGCTTGAAGGTTGAGAGACACTTCCATTTGAAGATACATGATGAGACT 896
DB 655 CTAGTTGAAGACATGATATGAGAGAGAGAGAGAGAGAGAGAGATCTTGTATCTT 714
QY 897 GCAACGAGAGCCGATGAT---CTTGTATCTGACGCTGATGAAGAAATACATCACTT 953
DB 715 GAGGATCATGAAGGATGATGATACCTGAAACCGTTGATGATGATGATGATGATGAT 774
QY 954 TATGGGCTGGAAGAGCCAGTCAATGATGCTGATGCTTGAAGGAGCCATGATGCTT 1013
DB 775 TATGAAACCCAGAAATCAATCTGATCTTCTGATGATGATGATGATGATGATGAT 826
QY 1014 GGAAGTGTGCTTCTTCTGATCTGTCATGAGAGATGTAAGAGATGATGATGATGAT 1073
DB 827 ATGCTCATCTTGGGCTTACGCTCTGCGCAGAGAGAGAGAGAGAGAGAGAGAGAG 885
QY 1074 ATGATCCGATGTCATCTTTTGTGATGTCATGATGATGATGATGATGATGATGAT 1133
DB 886 AAGATCCGCTGTCATCTTTTGTGATGTCATGATGATGATGATGATGATGATGATGAT 945
QY 1134 AGTATGAGACATGTTTCTTCAACCTTGAAGATGATGATGATGATGATGATGATGAT 1187
DB 946 AACCTGAGATGAGATTTTCTTCTGATGATGATGATGATGATGATGATGATGATGAT 1005
QY 1188 CAGATGCAAAATGAGCAGTGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1238
DB 1006 CTTACGATTAACCGGCTCATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1065
QY 1239 GAGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1289
DB 1066 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1125
QY 1290 AGCCATGCTGTCAGGAG 1349
DB 1126 AGCCCTTAAATGTCGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1182
QY 1350 CACGCTGAGATGCTTGTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1409
DB 1183 ACAATGAG 1239
QY 1410 GTGAGCAGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1469
DB 1240 GAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1279
QY 1470 GGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1529
DB 1280 -----ATGATATGAG 1329
QY 1530 CTTGCTCAAG 1589

Db 1330 GCTGAATCTCGCCGTCGATCATCTTATCCCGAGGTCGGATGCT-----GGA 1383
 QY 1590 AGTAGATTGTATACCTGCTGCTTGTAGTAAGTCAAGTCACTTTTCTCAAGGGCTT 1649
 Db 1384 GCGAGCTACATTACCGCTTCTGCTTGTAGCAGATCTTGTCTTGGCTTAATACAT 1443
 QY 1650 GCTGAACAAGCATGTCAAGTGTCCATGTTTCAACCATCTGAGTAGTCCAAAGT 1709
 Db 1444 CATGA-----TCGCGCATGTTCCCGGAGAAATGTGCTGCTCGGA 1488
 QY 1710 TCACCTGGAAAGATTGTGTGAACCTGAGTAGAGCGCTGCTGTATAGTCCGTGGA 1769
 Db 1489 CCACTCTGCTGCTCTTCTTGAACCTGAGTAAACGCTGCTGTGTGTGTGCTGCG 1548
 QY 1770 ATTACATCTTCAACAGTTTGTCTGAATTAACGAGTGTCTGTACTATACCCCAAAAT 1829
 Db 1549 ATTCAATATCTGAGAGATTTTCAAGTATCAATGAGTCTTCTACTACACTCTTCAGAT 1608
 QY 1830 CTTGACCAAGCTGTGTGCAATTATCTTCCAAATTTGCTCAGCTCGGCATGAGA 1889
 Db 1609 CTCGAACGGGCTGCGTAGATATCTTCTTCCAGCTCGGACTTAAGTTCATCTCGGG 1668
 QY 1890 TCATCTTGTATCACTTCTCTCACTACTATATGCTTCTTGTGCAATGCTTGTGCAATG 1949
 Db 1669 TCATCTCTCATCAAGCGGTTTAAACAATTAATCTATCTCCAGCCATGTGTGTGCAATG 1728
 QY 1950 CTGCTTATGATCTTCTCCGGAAGAAGTTTTGTCTGTAGGCAATTCATTCATCTTGATA 2009
 Db 1729 AGACTATATGATATATCCGAGAAAGTCAATTAATCTTCTGTGCAAAATCCAGTTCTCAT 1788
 QY 2010 GCATCTTAAGTATCTGCTGTGTGTGCAATTAATTAATTTGGATACACTAGCCCATGCT 2069
 Db 1789 GTCTCACTGTGCTCTTGTGTATCATCAGGAGCTCATCAATCAAGCAAGTGTGAAAGCA 1848
 QY 2070 TTGCTCTCAACCGTCACTGTATATCTTACTTCTGCTGCTGTATTTGGATTTGGCTCC 2129
 Db 1849 GCATCTTCAACAGTGTGTGTGTGTGTCTTCTGCTTCTTCTGTATGAGGTTACGGTCCC 1908
 QY 2130 ATCCCAACATTTTATGTGAGAGATCTTCCAAACGAGGTTGCGGCTCTGTATTTGCC 2189
 Db 1909 ATTCCAAACATCTCTGTGTGTGAATCTTCCAAACAGAGTCCGTGCTCTGTGATCGCC 1968
 QY 2190 ATTGTGCTTATACATCTGTGATCGAGATATCATGTCACCTTACAGCTTCTGTGATG 2249
 Db 1969 ATATGTGATATGCTCTTTTGTGATGAGACATATGTGACATCACTTCCCTGTCTC 2028
 QY 2250 CTGAATGCTATGATCTGCGGGGTGTTTTCAGCAATATGCAATGCTATGCTGTATTTCC 2309
 Db 2029 CTCAGCTCATCGACTAGTGTGTGTGTTCAGCATTTACGCTGCGGTTTGCCTTATCTCA 2088
 QY 2310 TTTGTGTGTCTTCTTAAAGTCCCTGAGCAAAAGGGATGCGCTTGTAGGTTATTTACC 2369
 Db 2089 TGGATTTGTGTAAAGAAAGTCCCGAGACTTAAAGGATGCTTTGGAAGTTATCA 2148
 QY 2370 GAATCTTTGCAAGTGTGCTC 2389
 Db 2149 GACTACTTTTGCCTTTGGAGC 2168

RESULT 12
 AR208571 1487 bp DNA linear PAT 20-JUN-2002
 LOCUS AR208571 Sequence 13 from patent US 6383776.
 DEFINITION AR208571
 ACCESSION AR208571
 VERSION AR208571.1 GI:21509759
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1487)
 AUTHORS Allen,S.M., Hitz,W.D., Kinney,A.J. and Tingey,S.V.
 TITLE Plant sugar transport proteins
 JOURNAL Patent: US 6383776-A 13 07-MAY-2002;

FEATURES
 source 1..1487
 BASE COUNT 355 a 320 c 373 g 439 t
 ORIGIN

Query Match 24.9%; Score 691.2; DB 6; Length 1487;
 Best Local Similarity 80.1%; Pred. No. 1.2e-170;
 Matches 852; Conservative 0; Mismatches 203; Indels 9; Gaps 3;

QY 1395 GAGGGTGGATGTGTGTGAGCAGCATGTATCTGTTGGGGATGTGCAAGTCTTGGAAA 1454
 Db 12 GAGGGTGGGGAGGCGAGTGCAGCAGCATGTATTTGGGGGGGTGGCAATCCCATGAAA 71
 QY 1455 TGGTCAAGAAAGAAAGTGTGAAATGTGAAAGAAAGTGTTCAAAAGATCTACTTG 1514
 Db 72 TGGTGGAGCGCAAGGCGAGATGTGCAAGAAAGAGGCTTCAAAAGAAATCTACTTG 131
 QY 1515 CACCAAGAGGAGTCTTGGCTCAAGAAAGGAGCTCAATTTTTCATCTTCCGTTGGGC 1574
 Db 132 CACCAAGAGGAGTGGCGAGCTCAAGAAAGGAGCTCTGTTTTCATCTTCCGTTGGGC 191
 QY 1575 GATG--TTTGAAGGTATGTAGATTGTGTACATGCTGCTTGTATGATGATGAGCA 1631
 Db 192 GATGCCAGCAAGGGGGAGTGGGTTTATACATGCTGCTTGTGATGACCATCGGCT 251
 QY 1632 CTTTCTCAAAAGGCTTGTCTTAAACGAGATGTCAAGTCT--GCCATGTTACCCA 1688
 Db 252 CTTTACTCCAAAGATCTTATGAAAGAGCTATGAGCGGCGGTCAGCCATGATTCACA 311
 QY 1689 TCTGAGGTAGTGTCCAAAGGTTCAAGTGTGAAAGATTTGTGAACTGTGAGAGGCT 1748
 Db 312 TTGAGAGCAGTCTCCAAAGTTCATCTGAAAGATCTGTTGAACTGTGTGAGGCT 371
 QY 1749 GCCCTGTATGTGTGTGTGAATTCAGATCTTCAACAGTGTGTGAAATTAACGCTGT 1808
 Db 372 GATTTGTGTGTGTGTGAATTCAGATCTTCAAGAGTGTGTGAAATTAATGAGAT 431
 QY 1809 CTGTACTATACCCCAAAATTTCTTGAAGAGTGTGTGAGGATTTCTTCCAAATTT 1868
 Db 432 CTGTACTATATCTTCAAAATTCGAGAGCAAGGTGTGTGTCTTCTTCCAAATTT 491
 QY 1869 GGTTCAGCTGGGCAATGATCCATCTTGTATGATCTTGTGATCTTGTGATCTTGTGAT 1928
 Db 492 GGTTCAGTTCAGTATGATCCATCTTGTATGATCTTGTGATCTTGTGATCTTGTGAT 551
 QY 1929 CTTTGATTTGCTTGTGCAATGCTGTATGATCTTGTGCAAGAAAGTGTGTGCTTA 1988
 Db 552 CCAAGATTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 611
 QY 1989 GGCACATTTCCAAATCTTGTATGATCTGTATGATCTGTGTGTGTGTGTGTGTGTGT 2048
 Db 612 GGCACATTTCCAAATCTTGTATGATCTTGTATGATCTTGTGTGTGTGTGTGTGTGT 671
 QY 2049 TTGGGTACATGAGCCATGCTTGTGTGCTTCAACGTCAGTGTATGCTACTTGTGCTGC 2108
 Db 672 TTGAGTAGGCTCCCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 731
 QY 2109 TTTGTATGAGATTGT 2168
 Db 732 TTTGTATGAGATTGT 791
 QY 2169 GTTGT 2228
 Db 792 GTTGT 851
 QY 2229 ACTTACAGCTTCTGT 2288
 Db 852 ACTTACAGCTTCTGT 911
 QY 2289 GAGATGTATGCTGTATTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2348
 Db 912 GAGATGTATGCTGTATTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 971

QY 2349 ATGCCCTTGGATTATTAACCAATTCCTTGCACTGGTGGCCAGCAAGC---GGTCGA 2405
DB 972 ATGCCCTTGGATTATTAACCAATTCCTTGCACTGGTGGCCAGCAAGCAGCCAC 1031
QY 2406 AAGACCTAATTCCTTGGTACCTTGGTGGCAACTATTCGACTG 2449
DB 1032 ATTGGCTGATTCATCATGAGAGCTTTGTTTCACTTGGACACTG 1075

RESULT 13
ATSGUTRPR
LOCUS A.thaliana mRNA for sugar transporter. linear PLN 17-AUG-1996
DEFINITION 2426 bp
ACCESSION 250752
VERSION 250752.1 GI:1495272
KEYWORDS sugar transporter.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 2426)
Tjaden, J. and Neuhaus, E.
A new sugar transport protein from Arabidopsis thaliana
Unpublished
2 (bases 1 to 2426)
Neuhaus, E.
Direct Submission
Submitted (09-AUG-1995) Neuhaus E., University of Osnabrueck,
Pflanzenphysiologie, Barbarastr. 11, Osnabrueck, Germany, 49069
Location/Qualifiers
1. 2426
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
37..2241
/codon_start=1
/product="sugar transporter"
/protein_id="CAA90628.1"
/db_xref="GI:1495273"
/db_xref="SPTREMBL:O96290"
/translation="MKGATLVLAATIGFPLQGMNATITAGMYINHDIANPTSVQ
LVNANSLGATVITTCGPIIDMLGRPVLLSSVWTFVGLIMLSPVYVLCARL
LNGFAGLAVTLVPVYISETAPPELRQQLNTLPQLFGSGAMFLSCWFTMLSLSPS
WPMALGVLSIPSLVFLTVFVLPESPRMLVSGRMDEKRVLCQCGREDYDLMAL
LYEGLDIGEKTEMLDVLTEHEDGDTLETVEDQILRYGHEHOSVLAPVEON
SLGSRSHRSLANOSMILKDPVNLFGSLHEKMPAGNTRSGIPPHGSMSTTAD
AHGKRAHWKXDIESHYNKDNDDYATDDAGDDSDNLRSLPMSRQTTSDKMDIP
HPTSGSTLSRRHSTLMQNGSSWIGGGMHGRYRDEYKRYLKEGDGSRGGS
IISIPGPDGGSYIHASALVSRVSGPVSGHSAVPEKLAASGPLMSALLEPQV
RALVVGVIQILQDFSGINLVYTPQILERAGVIDLSLSLSISAFLLSGATTL
LMLPAIVAMRLMDVSGRRSLMLMTPLVILVIVISLHISVVNVALSTGCIV
LYCFEPMGVPFOTSSVLKSSQADRGCLICAMVFWIGIITVYSLPVLSSEL
VGVPSIYAAVCVSIWTFVYMKYPERKGMLEVIYDYFAAGAAQASABSKDI"

BASE COUNT 577 a 569 c 544 g 735 t 1 others
ORIGIN

Query Match 24.2%; Score 672.2; DB 8; Length 2426;
Best Local Similarity 61.0%; Pred. No. 1.3e-165;
Matches 1567; Conservative 0; Mismatches 773; Indels 101; Gaps 13;

QY 177 ATGTGGGGGCTGTTCTTGTGCGCATATGCGCTCCATCGGCATTTATGCAAGGGGTG 236
DB 37 ATGAAGGAGGACGCTCTCGTTGCTCGCGCCACATCGGCATTTCTTCAAGAGATG 96
QY 237 GACATGCCACATGCCAGCTGCTGTTCTGATATATAAGAGAAATTTCAATTCGAAAT 296
DB 97 GACATGCCACATGCCAGCTGCTGATATATATACCAAGAACTTGAATCTACCA-- 154
QY 297 GAGCCCACTGTGAGAGGACATAATGTGTCAATGCTATTATGAGGCGCCACATCGTTACT 356
DB 155 --CTCTGTTCAAGGCTTTGTCGTTGCTATGTCAATTGATCGGTGCAACGGTCATCAG 210

QY 357 ACATTCCTCGGCGCATTAACAGATCAGATTGGCCGACGCCCTTATGCTTCTTCA 416
DB 211 ACTTCCTAGAGACCATATATGATTTGGCTCGGACAGCCCATATCTCATTTATATCA 270
QY 417 ATTCTGATCTTTAGGAGGCTCATGACATAGTGTCTCCAAATGTCATGATGCTG 476
DB 271 GTTATGATTTTGTCTGCGGTTTATTAATGTTGTGCTCCCAATGTCTATGTTCTG 330
QY 477 TTGGACGCTTCGTAGATGATTTTGGTATTTGGCTTGGCTGACGCTTGTGCTTAC 536
DB 331 TTGCTAGGCTTCTTAATAGGTTTGTGCGCGGCTCGGGTTACACTGTCCTGTTAC 390
QY 537 ATTTCGAATAATGCCCCCTTCGAGATTTAGGATTTGCTGATACATACCAATTCAGT 596
DB 391 ATTTCTGAACCGCTCTCCGAGATCAGAGACAGTTAAATACCTCCCTCAGATTTCT 450
QY 597 GGATCAGAGGAATTTCTGTATACAGCATGAGTTGGATGGATGCTGCGGCATCA 656
DB 451 GGCTGTGGATTTTGTGATCTGATGATGATGATGATGATGATGATGATGATGATG 510
QY 657 CCCGATGAGAAATTAATGCTTGTGCTGCGGATACCTTCATGTTCTTTGATTTG 716
DB 511 CCTAGCTGAGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 570
QY 717 ACAATATTTTATCTTCTGATCTTCGAATCTTCGAATGCTGCTGCTGCTGCTGCTG 776
DB 571 ACGGTGTTTATTTGCTCCGAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 630
QY 777 GCAAAAAGGCTGCAAAAGTTTACGAGGGAAGAGATGCTGAGTGAATTTGCCCTT 836
DB 631 GTTAAGCAGATTTCTTCAACATTTATGTCGACAGAGATGTTTACCATGATGATGCTT 690
QY 837 CTTCGAGAGGCTTGAAGCTTGAAGACACCTTCATTAAGATATCATCATCATTCAG 896
DB 691 CTAGTTGAAGACATGATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 750
QY 897 GCCACCGAGGACCGCATGAT---CTTGTACTACGCTGATAGAGAAATATCACACTT 953
DB 751 GAGGATCATGAAAGGATGATGATACCTTTGAACCGTTGATGAGAGAGAGAGAGAG 810
QY 954 TATGGGCTGGAAGAGGACGATGATGATGCTGACCTTTGAAGGAGCCCATCATGCTT 1013
DB 811 TATGGAACCCACAGAAATCAATCTGATCTGCTGATGATGCTGCTGCTGCTGCTGCT 862
QY 1014 GGAAGTGTCTTCTCTTCTGATCTGCTGATGAGAGACATGATGAGACAGATGACCTT 1073
DB 863 -ATAGCTACCTGGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 921
QY 1074 ATGATCCGATTTGACACTTTTGTGTGTGTGCTCATGAGATATGCTTCAGCTGAGA 1133
DB 922 AAGAGTCCGCTGCTCAATCTTTTGGCAGTCTCCACGAGAGATGCCAGAGACAGCGGA 981
QY 1134 AGTATGAGAGACATGTTTCCAAACTTTGGAATATGTTTCAAG-----TGTCAAGAT 1187
DB 982 AACACTCGAGAGGATTTTCCCTCATTTCCGAGCATGTTTCACTACTACTGCTGATG 1041
QY 1188 CAGCATGCCAAATATGACAGTGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1238
DB 1042 CCTCAGGTAACCGGCTCATTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1101
QY 1239 GAGGATGACATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1289
DB 1102 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1161
QY 1290 AGCCATGCTGCTCAGGAG 1349
DB 1162 AGCCCTTAATGTCGCGGACAGACAGAGATGAG---CAAGAGATGATGATGATGATG 1218
QY 1350 CACCGTGAAGTCTTTGAG 1409
DB 1219 ACAATGGAAGACCTTTAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1275

QY 1410 GTGAGCAGACTGATATCGTGGGAGTGCAGCTTGCTTGAAATGTCAGAGAGAA 1469
 DB 1276 GAAAGAGACGAGGAATGATGGTGGTGGCATATGGAAAT----- 1315
 QY 1470 GGTGAGATGTTGAAAGAGAGTGTCTTCAAAAGTCTACTTGCACCAAGAGGAGTT 1529
 DB 1316 -----ATGAGATACGAAAAAGATGATACMAAGGATATTTTAAAGAAATGGA 1365
 QY 1530 CTTGGCTCAAGAAAGGGGCTCAATTTGTTCACTTCCGGTGTGGCGAGCTTTTGAAGGT 1589
 DB 1366 GGTGATATCTCCGGTGGCTGGATCTCTATTTCCGGAGAGTCCGATGTT-----GGA 1419
 QY 1590 AGTGAATTTGATCAGTCTGCTGCTTGTAGTACAGTCAAGCACTTTTCAAGGCTCTT 1649
 DB 1420 GGCAGACTACATTCACGCTTCTGGCCCTTGAAGAGATCTGTTCTTGCTCTTAATCAGTT 1479
 QY 1650 GGTGAACCAAGCATGTCAGATGTCGATGTTCAACCCATCTGAGGTAGCTGCCAAAGT 1709
 DB 1480 CATGGA-----TCCGGCATGTGTTCCCGAGAAATTTGCTGCTCTGGA 1524
 QY 1710 TCAGCTTGAAGATTTGTTTGAACCTTGAAGAGGCGTCCCTGTAGTCCGGTGTGA 1769
 DB 1525 CCACTCTGATCTGCTTCTTGAACCTGAGTGAAGCGTGCCTTGTGTTGTGTCGAGC 1584
 QY 1770 ATTCAATCTTCAACAGTTTGTGTAATAAAGGTGTTCTGTACTATACCCCAAAAT 1829
 DB 1585 ATTCAATCTGAGAGAGTTTCAAGTATCAATGAGTTCTGTACTACACTCTCTGATTT 1644
 QY 1830 CTTGAGCAAGCTGTGTGAGTATTTCTTCCAAATTTGTTCTAGCTGGCAGTACGA 1889
 DB 1645 CTCGAAACGGGCTGGCTGATATTTCTTTCGAGCTCGAGTAAAGTTCCATCTCTGCG 1704
 QY 1890 TCCATCTTATGATCTCTGCTACTACTAATGCTCTTCTGATGCTTGGCTTGGCAG 1949
 DB 1705 TCAATCTCATACGCGGTTTAAACATTAATCTCATCTCCAGCCATGTGCTGTCAG 1764
 QY 1950 CTGCTTATGATCTTTCGGAAGAGTGTGCTGCTAGGACCAATTCCTTGATA 2009
 DB 1765 AAGCTATGATGATGATCCGGAAGAGTCAATTAATCTCTGAGCAATCCAGTTCATT 1824
 QY 2010 GCATCTTATGATCTTCTGCTGTTGTTCTCAATCTAATGATTTGGATACCTAGCCCATGCT 2069
 DB 1825 GTCTCACTTGTCTCTTGTGATCAGAGCTCATCATCAGCAAGTCTGGAACGA 1884
 QY 2070 TTGCTCTCAGCGTGTGATGCTCTCTCTGCTGCTGTTATGAGATTTGGTCCC 2129
 DB 1885 GCACTCTCAGAGTGTGTGCTGCTCTCTCTCTCTCTGATGAGGTACGAT-CC 1943
 QY 2130 ATCCCAACATTTTATGTCAGAGATCTTTCACACAG-GGTTCTGAGCTCTGATTCG 2188
 DB 1944 ATTCCAAACATCTCTGTTCTGAAATCTTCCCAAGAGAGACCGTGTCTCTGATTCG 2003
 QY 2189 CATTTGTGCTTATCAATCTGATGTCGAGATATCATGTCACCTACAGCCTTCTGTGAT 2248
 DB 2004 CATATGCTATGATCTTTTGTGATGAGACATTAATGTCAGTACTCTCCCTCTCT 2063
 QY 2249 GGTGATGCTATGATGTCGCGGCTTTTTCAGCATATATGAGTGTATGTTTC 2308
 DB 2064 CCTGAGCTGATGAGATGATGAGATTTTTCAGCATTTACGCTGCGGTTGTGCTATTC 2123
 QY 2309 CTTTGTGCTGCTTCTTAAAGTCCCTGAGCAAAAGGAGATGCCCTTGAAGGTTATTC 2368
 DB 2124 ATGATCTTCTGTTATGATGAAATCCCGAGACTAAAGCATGCTTTGGAAGTTATCAC 2183
 QY 2369 CGAATCTTTGAGTGTGTC 2389
 DB 2184 AGACTACTTGTGCTTGGAGC 2204

RESULT 14
 ATH532571 2190 bp mRNA linear PLN 11-DEC-2002
 LOCUS
 DEFINITION Arabidopsis thaliana mRNA for monosaccharide sensing protein 3

ACCESSION (mesp3 gene).
 AJ532571
 VERSION AJ532571.1 GI:26800696
 KEYWORDS monosaccharide sensing protein 3; mesp3 gene.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1
 Stamme, C., Tjaden, J., Trentmann, O., Emmerlich, V. and Neuhaus, E.
 A novel family of plant monosaccharide transporters is involved in a new type of eukaryotic sugar sensing
 REFERENCE 2 (bases 1 to 2190)
 Neuhaus, E.
 Direct Submission
 Submitted (04-DEC-2002) Neuhaus E., Plant Physiology, University of Kaiserslautern, Erwin-Schroedinger-Str. 22, 67663 Kaiserslautern, GERMANY
 FEATURES
 source location/Qualifiers
 1..2190
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 1..2190
 /gene="mesp3"
 1..2190
 /gene="mesp3"
 /codon_start=1
 /product="monosaccharide sensing protein 3"
 /protein_id="CAD58693.1"
 /db_xref="gi:26800697"
 /translation="MRSVTVLAAALNMGDMNATAGAVIYIKKPHLEKPKI
 EGLIVAMSLIATLITPFSVPDVGGRSLILSSVYFSLSIYMFSPVYVLLFA
 RLIDFGILAVTVLPVYISBTAPELRLNTPQFCGSGMFLSYCLVGMGLQES
 PSMRLMLGLVLSIPSLAYFLVLAFLPEPRMLVSGKMDKAROVLRGRGDEYSEL
 ALLVGLGVGKDTSEIEYVIGPDNENSGNELPKKDOIKLGPEDGSMMAKPVKGO
 SIALASROGSMTPRGSGIMDPLVTLRPSIHEINLPSRMNASSSMTPEPMNGSLIGM
 GROSQMDPERNNEDSSPDENLNSPLSPOTTERPDYHQRTVGTMRGOSLFRANV
 GETATATISIGGQWQIAMRYNDKVGADGRVNGLQRMVTHEETANNNTNNTIPSRGS
 LLSFPEBDGHDVNGVQAALVSOASMGGKGETMLPKRYVDGGMELKRPV
 KRALMVGVLQILQOPAGINVMYTPQILEETGVSLLTMGISASBASLLISALT
 LLMLPCLISMRSLMLSTIPILISLVTLVGLSVNLGGGSLNALISLTSVYVLSFV
 MGRGAPILILCSHLEPSSVRLCTICATLFWICDIIVTTLTPMLKSLIGAGVGIY
 ALVCVAVNFYLLKVPETKMPLEVISSEFVSVAQDDAASPLSDG"
 BASE COUNT 549 a 476 c 569 g 596 t
 ORIGIN
 Query Match 23.3%; Score 646.2; DB 8; Length 2190;
 Best Local Similarity 58.6%; Pred. No. 9.2e-159;
 Matches 1321; Conservative 0; Mismatches 848; Indels 86; Gaps 8;
 QY 177 ATGCGGGGGCTGTCTTGTGCGCATAGTCGCTCATCGGCAATCTATTGAGGGGCTGG 236
 DB 1 ATGAGAGAGTATGCTCTTGTGCTTACCGCGTGAGATAGGAATATGTTGAGGGCTGG 60
 QY 237 GACAAATGCCCATGCGAGCTGCTGTTCTGATATTAAGAAAGAAATTTCAATTC 296
 DB 61 GACAAATGCCCATGCGAGCTGCTGATTAATTAAGAAAGAAATTTCAATTTGAGAGAA 120
 QY 297 GAGGCCACGTGAGAGGAGCTAATTTGTCATAGTCACCTTAATGCGGCCACCATGCTTACT 356
 DB 121 GAACCAAGATAGAAAGACTGATCTGGCAATGCTCTCATATGAGGCCACTTGAATACCA 180
 QY 357 ACATTTCTCGGGCCATTAATGAGCTGATTTGGCCAGACCCCTATGCTTATCTCTTCA 416
 DB 181 ACTTCTGCTGCTGCTGCTGCTGAGCAAGATAGGAAGGAGATCAATGCTTACTCTCTCT 240
 QY 417 ATTCTGACTTTTGTAGGGGCTCATCATGCTATGCTCTCTAATGCTATGCTGCTG 476
 DB 241 GTTCTCTATTTTCTGATGATGATGTTATGTTTGTGCTCCCAATGCTACGTTCTCTCT 300

Oy	477	TTGGACACCTTCGTAGATAGATATTTGGTATTTGGCTTGGCTGTCACACTTGTGCTTTGAC	536
Db	301	TTGCGAAGGCTTCTTGATGTTTGGATCGGTTTAGCTGTCACTCTGTGCCAATCTAC	360
Oy	537	ATTTCAGAAATAGCCCTTCGGAGATTAGAGTTTGTCTGAATACCTACCAAAATTCAGT	596
Db	361	ATCTCTGAGACCGGCACTTCTGAGATCAGAGGATTAATCAACACTTCCCGAGTTTGT	420
Oy	557	GGATCAGAGAGAAATGTTCTTGTCAATCTGCATGCTGTGGATGTCCCTGTCCGATCA	656
Db	421	GGATCCGGTGGGATTTTTTGTGCTAATGTCTTGTGTTGGGATGTCCCTTCAMAAATCA	480
Oy	657	CCCGATTTGAGAAATATATGCTTGGTGGTCCGGATACCTTCATTTCTTCTTGGTTG	716
Db	481	CCACCTGAGGCTATGCTTGTGTGTGTGTGTCTTAATTCGTCATTTGCTTACTTGTACTT	540
Oy	717	ACAAATATTTATTTCTTCTGAATCTCCAAAGTGGCTGTAGCAAAAGTCGATGACAG	776
Db	541	GGCGCTTTCTTCTTCCCTGAATCTCCAAAGTGGCTGTAGCAAAAGTCGATGATGATA	600
Oy	777	GCMAAAAGGTGTCTCCAAAAGTTACGGGGGAAAGCAGTGTCTCAGGTGAATGTCCCTT	836
Db	601	GCTACACAGTCTCTCAGAGACTCCGTGGCAGAGAAAGATGTTTCAGGGAGCTTGTCTG	660
Oy	837	CTTCTCGAAGGTTGAGGTTGGAGGAGAGACATTCATTTGAAGATACATCATCTTGGACT	896
Db	661	CTGTGTGAAGGCTTGTGGGTGGGAAAGACAGTGCATTAAGAAATATGTGATGGACCA	720
Oy	897	GCCACCGAGGCGCCAGCTGATCTTGTACTGACGGTGAATAGGAACAAATCACTTAT	956
Db	721	GAGACGAGGAAAGAGAGGTGGAAATGAACGTCCGAGAAAGATCAAGTAAACTATAT	780
Oy	957	GGGCTTGAAGAGCCAGTCATGATTTGCTGACCTTCTAACGGACCATCATGCTTGA	1016
Db	781	GGTCCAGAGATGAGCAGTCATGATGCTTAAGCCAGTGAAGAGAC-----AG	828
Oy	1017	AGTGTGCTTTCTCTTGAACTCTGATCTGATGGGAGACATGTGAACCAAGTGTACCCCTTATG	1076
Db	829	AGTTCTTCTCGCAATTTGAGCTTCCGACAAAGCCAGCATTTACCGGTGGCGAATCCCTCATG	888
Oy	1077	GATCCGATTTGTGACATTTTGTGTGTGTGTCATGTGAATATATGCT-----CAAGCTGA	1138
Db	889	GACCCACTTGTCACTCTCTTGTGTGACATTCATGAAATCTCCCTCCGAGAAACATGAC	948
Oy	1131	GGAAATATGAGAGACATTTGTTTCAAACTTTGAAAGTATGTTCAGTGTCAACAGATCAG	1190
Db	949	GCATCATCCCGCAGATGCTCTTCCCAATATGGAGATATCTGGAGATATGAGGAGG	1008
Oy	1131	CATGCCAAAATATGACGATGGGATGAAGAAATCTTCATAGGAGATGACAGAGATACGA	1250
Db	1009	CAGGATGCCCATGTGGATTCACAGAGAAACATATGAAGTATGTTCTGAACCAATATAAAA	1068
Oy	1251	TCGTATGTTGTCAGAGAGGTGACTATATAGGACAAATCTCCATAGCCCATTCGTGTCAGGCGAG	1310
Db	1069	TTTAAA-----CAGTCTCTGCTTTTCTCCGAAAACCATGAGCCGATGACTACACCAAGC	1122
Oy	1311	GCACAGGTCCGAAAGGAGGAGACATTTGTGACATGTGCACCGTGAAGTGTCTTTGAGC	1370
Db	1124	GGACCGTTGTATCANTGACATAGGCGACAGAGCAGCTGTTTATATGCAAAACGTGGGTGAGA	1183
Oy	1371	ATGAGAAAGCAAAACCTCTTAGGGAGGGTGGAGATGTGTGTAGACAGCACTGATATCGGT	1433
Db	1184	CAGCCACGGCTTACAGACATAGCGGTGATGGCAATTTGCCGTGAAAGTATA-----	1233
Oy	1431	GGGGGATGGCAGCTTGTGAAATAGTTCAGAGAAAGAA--GGTGAAGATGTGAGAAAGGA	1488
Db	1235	-----ATGACACAGTTTGTGTGCAATGTATAGAGAGTCATATGAGGGTTTACAGAAATGTA	1288
Oy	1490	AGGTGATTTCAAAAGATCTTATCTTGACACCAAGAGGAGTTCTCTGCTCAAGAAAGGGCTC	1548
Db	1290	TATTCACGAGAAACCGCCAAACAAACACCAACAAATCTCTTTTGGCGACGTGTGATC	1349
Oy	1550	AAATGTTTCACTTCCCGGTGTGGCGAGTGTTTTG---AGGATGATGATTTGTATCATGTC	1606

Db	1350	ACTTCTCTCTCCACCCACGAGGTGATGGTCAATGATCAGGTGAATGATACGTTACGGC	1409
Oy	1607	TGCTGCTTTAGTAAGTCAGTACGACTTTTCTCAAAAGGCTTGTCTGAACCAACGCAATGTC	1666
Db	1410	TGCTGCACCTTGTGAGCCACACTTTCGATGATGCGCAGAGGTAAAGGCCAGACCCGCAATGT -	1468
Oy	1667	AGATGCTGCCATGGTTACACCATCTGAGGTGATAGCTGCCAAAGGTTCAAGTTGGAAGAATTT	1726
Db	1469	-----TGCCAAAGAGAGTTAAAGAAATGCTCAGGCTGAGAGGAGCT	1508
Oy	1727	GTTTGAACCTGGAATGAGCGCGTCCCTGTTAGTGCATGCTGTTGAAATTCATGATCCCTTCAACA	1786
Db	1509	GAAAGAACCAAGGGGTTTAAGAGACGTTTGAATGTTTGAAGTTGGGCTTCAGATACCTGCACAA	1568
Oy	1787	GTTTGTGGAATAAACGGTGTCTGTAATAACCCCAAAATTTCTGAGCAAGCTGCTGT	1846
Db	1569	GTTTGCAGGAATAAATGAGATGATGATTAATACCACTCAAAATATGGAAGAAACAGGTGT	1628
Oy	1847	GGCAGTATATCTTCCCAAAATTTGGTCTCAGCTGGGCAATGACATCCATCTTGAATCAGTTG	1906
Db	1629	GTCAAAGCTTTTGAACCAACCTTGGAAATGAGCAGAGCTGCAATGCGCTTCTCATAAACGC	1688
Oy	1907	TCTCACTACCTTAATAATGCTTCCCTGCAATTTGGCTTGCACATGCTGCTTAATGATCTTTTC	1966
Db	1689	CTTAAACCACTCTTGAATGCTTCCCTGCAATTTCTGTCTCAGT-----	1731
Oy	1967	CGAAGAGGTTTTTGTGCTAGGACACAATTTCCAAATCTTGAATAGACATCTTGAATTCCT	2026
Db	1732	-----AGGCTCTGATGCTTTGCACTATCCCATCTTAATCTGTGCTGATGAACAT	1784
Oy	2027	GGTGTGTCCAAATTAATTTGTTGGGTACATGACCAATGCTTTGCTTCCACCCCTCAG	2086
Db	1785	GGTATATGGAAGCTTGAATCTTGAAGGTTCAATTAACGCGTGAATATGACAGACAAG	1844
Oy	2087	TGTTATGCTCACTTCTGCTGCTCGTTATGGAATTTGTCCTCCATCCCAACATTTTATG	2146
Db	1845	TGTTAACGTTGACTTAAGCTGTGTTTCGTGATGGGTTTTGGGGCAATTCCAACATCTCTG	1904
Oy	2147	TGCAGAGATCTTTCCAAACAGGGTTCGTGGCCTCTGTATTTGCCAATTTGTGCTTTACAT	2206
Db	1905	CTCAGAGATATATCCCTCACTTCTGTGGCGCGGCTCTGCATCAACCAATGTGGCCCTCACAT	1964
Oy	2207	CTGATATGGAATATATATCTGCTCACTTACAGCTTCTCTGTATGCTGAATGATATTTGAAT	2266
Db	1965	CTGATATCTGTGACATTAATCTGCTCACTTCTCCACGATCACTGCTCAAAATTCATTTGCAT	2024
Oy	2267	GGCGGGTGTTTTGACATATATGACAGTCGATGCTTGAATTTCTTTGTTGCTGCTCTCT	2326
Db	2025	CGCAGAGATCTTTGGAATTAATGCAATGCTGTGCTGTGCGGTGGTTTTTGTGTACT	2084
Oy	2327	TAAAGTCCCTGAGACAAAGGGGATGCCCCCTTGAAGTTATTAACGAATCTTTTGCAGTTGG	2386
Db	2085	CAAGGTACAGAGACAAAGGGGAATGCCCTTTGAAGTTATCTCTGATGTTCTTCTCGTCGG	2144
Oy	2387	TGCGAAGCAAGCGGCTGCCAAAAGCTTAATTTCTTT	2421
Db	2145	TGCAAAACAGCAAGACGCTGCAAGCTTCAATTTCTCT	2179
RESULT 15			
AC136843			
LOCUS	AC136843	135406 bp	DNA linear
DEFINITION	Oryza sativa (japonica cultivar-group)	chromosome 11	clone
ACCESSION	OSUNBA0094M06, complete sequence.		
VERSION	AC136843		
KEYWORDS	AC136843.4	GI:31442544	
ORGANISM	HTG.		
ORGANISM	Oryza sativa (japonica cultivar-group)		
ORGANISM	Oryza sativa (japonica cultivar-group)		
ORGANISM	Eukaryota: Viridiplantae; Streptophyta: Embryophyta; Tracheophyta;		
ORGANISM	Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae;		
ORGANISM	Ethnatoideae; Oryzaceae; Oryza.		

REFERENCE 1 (bases 1 to 135406)
 AUTHORS Bhargava,A., Dalal,V., Sharma,T.R., Pal,A.K., Singh,A., Yadav,M.,
 Gaikwad,K., Srivastava,S., Dixit,A., Ghazi,I.A., Batra,K.,
 Mohapatra,T. and Singh,N.K.
 TITLE Complete sequence for Oryza sativa chromosome 11 Clone
 OSUNBa0094M06
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 135406)
 AUTHORS Ghazi,I.A., Singh,A., Srivastava,S., Dixit,A., Bhargava,A.,
 Pal,A.K., Dalal,V., Batra,K., Sharma,T.R., Mohapatra,T. and
 Singh,N.K.
 TITLE Direct Submission
 JOURNAL Submitted (08-NOV-2002) IIRGS, NRC on Plant Biotechnology, Indian
 Agricultural Research Institute, IBS Centre, New Delhi, Delhi
 110012, India
 REFERENCE 3 (bases 1 to 135406)
 AUTHORS Swain,S.C., Gaikwad,K., Pal,S., Yadav,M., Sureshabu,K.,
 Ghazi,I.A., Singh,A., Srivastava,S., Dixit,A., Bhargava,A.,
 Pal,A.K., Dalal,V., Batra,K., Sharma,T.R., Mohapatra,T. and
 Singh,N.K.
 TITLE Direct Submission
 JOURNAL Submitted (28-FEB-2003) IIRGS, NRC on Plant Biotechnology, Indian
 Agricultural Research Institute, IBS Centre, New Delhi, Delhi
 110012, India
 REFERENCE 4 (bases 1 to 135406)
 AUTHORS Bhargava,A., Dalal,V., Sharma,T.R., Pal,A.K., Singh,A., Yadav,M.,
 Gaikwad,K., Srivastava,S., Dixit,A., Ghazi,I.A., Batra,K.,
 Mohapatra,T. and Singh,N.K.
 TITLE Direct Submission
 JOURNAL Submitted (06-JUN-2003) IIRGS, NRC on Plant Biotechnology, Indian
 Agricultural Research Institute, IBS Centre, New Delhi, Delhi
 110012, India
 COMMENT On Jun 6, 2003 this sequence version replaced gi:28604240.
 FEATURES
 source
 1. 135406
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /db_xref="taxon:39947"
 /chromosome="11"
 /clone="OSUNBa0094M06"
 BASE COUNT 39875 a 28472 c 27393 g 39665 t 1 others
 ORIGIN
 Query Match 23.0%; Score 637.8; DB 8; Length 135406;
 Best Local Similarity 56.6%; Pred. No. 2.1e-156;
 Matches 1327; Conservative 0; Mismatches 972; Indels 46; Gaps 6;
 QY 84 CCGTCAGTGTCTTCTGCTTCTGCGGCTCTTCCGCTCTTCGATTCCTGACAGAG 143
 DB 99067 CTGTGGTTCGACAAATCATCTCAATCCCTCTGCGCTTTTCATGTGGTTTCTTGAGA 99126
 QY 144 CCGGTGACCTCGAGCATATCTTGGAGACAGATGTCGGGGCTGTTCTTGTGCGCAT 203
 DB 99127 AACCGATCCGAGATCGATCGATCGCCACGTAAGTAGAGGCGCGGTGGTGCGTG 99186
 QY 204 GTGGCTTCATCGGCATCTATTGCGAGGGTGGAGACATGCCATCGAGCTGCTT 263
 DB 99187 GCGGCGGCATCGGCAATCACTGCGAGGATGGAGACACGCACTCGCGCGCGCGTG 99246
 QY 264 CTGTATATAAAGAGATTTCAATTGCAATGAGCCCACTGTGAGGGAGCAATTTG 323
 DB 99247 CTGTACATCAAGGAGAGTTGCGCTCGAGACCCAGCGCGCTGAGGGCTGCTGTC 99306
 QY 324 TCAATGCTTATCGGCGGCACCATGTTACTACATCTTCGGGCGCATATGACATCG 383
 DB 99307 GCCATGCTCTCAATCGGCGGCACCATGATCAACCACTTCGCGCGCGCTCGACATCG 99366
 QY 384 ATTGGCGGAGCGCTTATGCTTCTTCTTCAATTTGTACTTTCTTACGGGCTTCATC 443
 DB 99367 GTGGGCGCGCGCGCATGCTATGCTCTCTCTCTTCACTTTCGCGCGCGCTTCATC 99426
 QY 444 ATGCTATGCTCTCTATGCTATGCTCTGCTGTGGACGCTTCGATGATTTGTT 503

DB 99427 ATGCTGTGTCCTCCCAAGCTACGCTCTCTCTGCGCGCTGTGACGAGCTTCGAC 99486
 QY 504 ATTGGCTTGGCTGTACAGCTTGTGCTTCTTGTACATTTGCAAAATGCCCCCTTGAGATT 563
 DB 99487 GTGGCGCTCGCGTCACTGCTCTCTCCGCTTACATCTCGAGACTCCCGCGGAGATC 99546
 QY 564 AGAGGTTGCTGAATACACTACCAATTCAGTGATGAGAGATGTTCTTGCATAC 623
 DB 99547 CGGCGCGGCTCAACAGCTGCGGAGAGTTCAACCGGCTCCGGGAGATGATGCTCAG 99606
 QY 624 TGCATGATCTTGGAGATGCTCTGCGCATCACCCGATGGAATGGAATTAATGCTGTG 683
 DB 99607 TGCATGATCTTGGAGATGCTCTGCGCATCACCCGATGGAATGGAATTAATGCTGTG 99666
 QY 684 CTCGGGATACCTTCAATGTTCTTCTTGTGTTGAACAATTTATCTTCGATCTCA 743
 DB 99667 CTCTTGTCTCCCTGCTCTGCTTCACTGTTGTCACCGCTTCTTACCTCCGAGTCCGG 99726
 QY 744 AGATGCTCTGTAGCAAAAGTGTGATGAGGAGGCAAAAAGGTTGCAAAATGTTACGG 803
 DB 99727 CGGTGCTCTGTAGCAAAAGTGTGATGAGGAGGCAAAAAGGTTGCTGATGCTGCGC 99786
 QY 804 GGGAAAGACATGCTCAAGTGAATTTGCTCTTCTTCTGAAAGGTTGAGGTTGAGAGA 863
 DB 99787 GGGCGGAGAGAGTCTCCGCGGAGATGGGCTGCTGTCGAGGGCTCGGACCGCGCGC 99846
 QY 864 GACACTTCTTGAAGATGATCATTTGAGCCTGCGACCGAGGCGAGCATGATCTTGT 923
 DB 99847 GACCCGAGATGAGAGATCACTGTCGCGCGCTCGAGGGGAGCGCGGCA-----G 99900
 QY 924 ACTGACGCTGTAGAAACAAATCACTTTATGAGGCTGGAAGAGGCGAGCATGAT 983
 DB 99901 AACGACAGGCGAGAGACACCTCACTGTCGAGGCGGAGACAGGCTTTCTGTGGTG 99960
 QY 984 GCTGACCTTCTTAAGAGACCCATCATGCTTGAAGTGTCTTCTTGCATCTGCTAT 1043
 DB 99961 GCGGACCGGTGCGCGCGCGCGCGGCGGAGCATGCTGGGAGCTCGTGGGCTGACAGGGC 100020
 QY 1044 GGGAGCATGTGAACCAAGATGATCCCTTATGATTCGATTTGACACTTTTGTGAT 1103
 DB 100021 TCGCGCATGAGCATGTATCAAGACAGATGAAGACCCCGTGTGGCGCTTCTCGGAGC 100080
 QY 1104 GTCCATGAGAAATATGCTCAAGCTGAG-----GAAGTATGAGAGCA 1148
 DB 10081 GTCCAGAGCGGCTGCGAGTCCGCGCGCGCGGCGGACCGGACATGAGGGAGACAG 100140
 QY 1149 TGTGTTCCAACTTTGAAATGTTCAGTGTCACAGATCAGC----- 1191
 DB 100141 CTGTTCCCAACCTCGGAGCATGCTTGAAGCTCAAGATAGCCCGGCGGAGCATGAG 100200
 QY 1192 -ATGCCAAATATGACATGAGATGAGAAATCTTCATAGGATGACGAGATGACGA 1250
 DB 100201 GACGAGAGAAAGTCAAGCTGCGACGACGACCTCGACAGAGAGAGAGAGATCTTC 100260
 QY 1251 TCTGATGTGACAGAGTGAATATGAGACATCTCCATAGCCCATTTGTGTCAGGAG 1310
 DB 100261 TCGACGACGCAAGAGACGACGACGAGCGCGGCGCTGACGACCGCTGTGTGCGCG 100320
 QY 1311 GCAACAGGTGCGGAGAGAGACATTTGACATGTCACCTGTGGAAGTCTTTGAGC 1370
 DB 100321 CAGA---GCACGAGTGTGAGACCAAGAACGAGCCGCTCGGCGAGGTGCGATGCG 100377
 QY 1371 ATGAGAGGCAAACTCTTACGAGAGGATGAGATGCTGTGAGACACATATTCGCT 1430
 DB 100378 CGGACAGAGCATGCGCGCGCGCGCGCTGAGACGCGGACGACCATGAGGATTCGGC 100437
 QY 1431 GGGGATGACGCTTGTGAAATGTCAGAGAGAAAGTGAATGCTGTAAGAGAA 1490
 DB 100438 GGGGGTGCACCTGGCGTGAAGTGAAGGAAAGTGGGCGCGGACGCGCTCAAGCGC 100497
 QY 1491 GGTGTTTCAAAAGAGTACTTATGACCAAGAGGAGTTCTGTGCTCAAGAGGAGCTCA 1550

Db 100498 GGCAGCGTGAAGGCATGTAACCTGACAGAGAGTCCAGGCCGCCGCCGGTGAACCTCG 100557
QY 1551 ATTGTTCA--CTTCCCGGTGTGCGGATGTTTTGAGGGTAGTGAATTGTAACATGCTG 1608
Db 100558 GCGCGCGCGCGGACCGGACGATGTCACGGGTAAGTCCACGCGCGCGGCTGTGAGCCGG 100617
QY 1609 CTGCTTTAGTAAGTCAAGTCAAGCACTTTTTCMAAAGGCTGTGCTGAACACGCAATGTCAG 1668
Db 100618 TCGATGCTGTACACCAAGGACGTCTGATCGGCGAGAGCCCAACGAGCCGGCGTTCGG 100677
QY 1669 ATGCTGCCANAGTTCAACCATCTGAGTAGCTGCAAGGTTCAAGTTGAAAGATTGTG 1728
Db 100678 AACCGGCC--GAGAGGGGTGGCGCGCGCGCGTGAAGGCCCGCGGTGGCGAGACTGC 100735
QY 1729 TTGACCTTGAAGTGAAGCGCTGCCCTGTAGTGGGTGTGGAATTCAGATCCTTCAACAGT 1788
Db 100736 TCGAGCGGGGCGTCCGCGCAAGCTCTTCTGCGGCGGTCAACATCCAGATCTTCAGCAAT 100795
QY 1789 TTGCTGGAAATMAACGGGTGTTGTGTAATATACCCCAAAATCTTGAAGCAAGCTGTGTGG 1848
Db 100796 TCTCCGGCATCAACGGCGGTGCTCTACTACACCCCGAGATCTCGACAGGCCGCGCTCA 100855
QY 1849 CAGTTATCTTTTCCAAATTTGTCTCAGCTCGGCATCAGCATCCATCTTGATCAGTTCTC 1908
Db 100856 GCGTCTCTCCGCGACGCTCGGCTCTCCGGCGACCTCACTCATCTCATAGCGGCC 100915
QY 1909 TCACATCTTACTTAATGCTTCTCTTGCATTTGCTTTCACATGCTGTATGATCTTTCCG 1968
Db 100916 TCACCAAGCTCTCCATGATCTCCGTCATGCGGCGTCCGCAATGCGCTCATGAGCGCTCG 100975
QY 1969 GAAGAAGTTTTGTGCTAGAGCAAAATCCAACTTGATAGCATCTAGTATGCTG 2028
Db 100976 GCGCGCGCGCTCTCTCTGTGAGCGCTGCGCTCTGCTGCGCTGCGCTGCTCG 101035
QY 2029 TTGTGTCAATCAATTAATGATTTGAGTACAATAAGCCATGCTTGTCTCCACCGTCAATG 2088
Db 101036 TGTGTGCGAAGCTGTGCGCATGCGCGGACCGCGACCGCGGCGCTGTGACGCGGAGCG 101095
QY 2089 TTATGCTACTTCTGCTGCTTGTGTATGAGGATTTGATCCCATCCCAACATTTATGTG 2148
Db 101096 TCATGCTTACTTCTGCTGCTGTGATGAGGATTCGCGCCCATCCCAACATCTCTGCG 101155
QY 2149 CAGAGATCTTTCCACAGAGGTCGTGCGCTCTGTATGCGCATTTGCGCTTACATTC 2208
Db 101156 CCGAGATCTTCCGACGAGGGTGAAGGAGCTGTGCAATGCGCATGTCTGCTGACGTTCT 101215
QY 2209 GGAATCGAGATATCATGTCACCTACAGCCTTCTGTGATGCTGAATGCTATTGACTG 2268
Db 101216 GGTGTGCGCATGCGCGTCAAGCTTCCGTCATGCTCAGCTCGTGGGCTCG 101275
QY 2269 CCGGTGTTTTCAACATATATGACGTGATGCTGTGATTTCTTGTGTGCTTCTTA 2328
Db 101276 CCGGCGTGTCTCTCTTCTAAGCGCGGTGTGCTGCGGCTCGTGTGTTGTTGCGCTCA 101335
QY 2329 AGGTCCCTGAGCAAGGGAGTCCCTTGAAGTTATTAACGAATCTTTGCACTGTGTG 2388
Db 101336 AGGTGCGCAGACCAAGGCGCTCCGCTGAGGTTCATCATGAGTTCTTCAACGTGCGG 101395
QY 2389 CGAAG 2393
Db 101396 CCAAG 101400

Search completed: January 5, 2004, 14:01:37
Job time : 1018 secs

XX Allen SM, Hitz WD, Kinney AJ, Tingey SV,
PI WPI, 2002-451386/48.
DR P-PSDB; AAU97201.
XX
PT New nucleic acid encoding plant sugar-transport proteins, useful for
preparing transgenic plants with altered carbohydrate distribution -
XX
PS Claim 3; Column 25-28; 54pp; English.
XX
CC The present invention relates to the isolation of plant polynucleotide
sequences encoding an Arabidopsis thaliana-like sugar transport
CC protein or Beta vulgaris-like sugar transport protein. The
CC polynucleotide sequences are useful for altering the level of sugar
transport protein in plants, i.e. for control of carbohydrate transport
CC and distribution in plant cells, e.g. during grain filling of annual
CC field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
CC carbohydrate flows and sugar transport. The polynucleotide
CC sequences can also be used to isolate cDNA sequences and genes that
CC encode homologues of the new proteins. The present sequence
CC representing a config assembled from various corn cDNA clones encodes
CC an Arabidopsis thaliana-like sugar transport protein.
SQ Sequence 2824 BP; 644 A; 649 C; 745 G; 775 T; 11 other:

Query Match 79.2%; Score 2200.2; DB 24; Length 2824;
Best Local Similarity 90.2%; Pred. No. 0;
Matches 2372; Conservative 0; Mismatches 253; Indels 6; Gaps 2;

QY 133 CTTCACAGAGCGCGTGACCTCGAGCATCTTGAGAGCAAGATGCGGGCGTTC 192
DB 194 CTTCGCGCGCGCTGCGGAGTGGGCGGTGATTTCCGCGGCGCATGGGGGCGCGTGA 253
QY 193 TTGTGCGCATAGTGCCTCCATCGGCAATCTATTGCAAGGGTGGAGCAATGCCACCATG 252
DB 254 TGGTCCCATCGGGCGCTTATCGGCAACTTGCTGCAAGGGCTGGCAATGGCAATGG 313
QY 253 CAGCTGCTGTTCTGTATATAAGAGAAATTTCAATTGCAAAATGAGCCCATGTGAGAG 312
DB 314 CTGGAGCCCGTCCGTATTAAGAGAAATTTCAACTGCGAGCGAGCCCTGATGAGAG 373
QY 313 GACTAATTTGTGAATGATCACTTATCGGCGGCGCAATGTTACTAATTCCTCGGGCAT 372
DB 374 GCTTCACTGTCGCGCATGTTCTCATTTGGGCGCAAGTATCAACAATCTCCGGGCGCA 433
QY 373 TATCAGACTCGATGCGCGAGCGCCATATGTTCTTCTTCAATTCGTACTTCTTCA 432
DB 434 GGGCTGACTGCGTTGGAGAGGCGCATGCTGCTGCGCTGCTGCTTACTTCTGCA 493
QY 433 GCGGCTCATCATGCTATGCTTCTTAATGCTTATGCTGCTGTTGGCAAGCTTGTAG 492
DB 494 GTGGGCTGTGATGCTTTGGGCGCGCAATGTGTACATCTTGCTCCCGAAGGCTATTG 553
QY 493 ATGATTTGGTATGCTGCTGCTGCTGCAAGCTTGGCTTTGTATCAATTTAGAAATGACC 552
DB 554 ATGGGTTGCGATGCTGTTGGCGGTACACTTGTCTCTTCAATCTCGAAATGCTGAC 613
QY 553 C-+-TTGAGAGATTAGAGTTTGGCTGAATACACTACCAATTCAGTGG-ATCAGAG 606
DB 614 CGAGACAGAAATTTGGGGGCTGNTGAAACGTTGCGAGTTCATTTGGGGTCAAGAG 673
QY 607 GAATGTTCTTGTATATCTGATGCTGTTGGAGTTCCTGCTGCGCATCAACCGATTGA 666
DB 674 GGAATGTTCTTCTCTCATGCTGATGCTTGGAGTTCCTCATGCGCCAAACGATTGA 733
QY 667 GAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 726
DB 734 GGGTCAAGCTTGAAGTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 793
QY 727 ATCTTCTGAATCTCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 786
DB 794 ACTGCTGATATCAAGAGTGGCTTGTAGCAAAAGAGAGATGGGAGGAGAGAG 853

QY 787 TGTTCGAAAGTTACGGGGGAAAGAGATGCTCAGTGAATTTGCTTCTTCGAG 846
DB 854 TGTTCGAAAGTTACGGGGGAAAGAGATGCTCAGTGAATTTGCTTCTTCGAG 913
QY 847 GGTTCGAGGTTGAGAGAGACTTCCATTTGAAGATATCATCTTGAAGCTTGAAG 906
DB 914 GTTTCGAGGTTGAGAGAGACTTCCATTTGAAGATATCATCTTGAAGCTTGAAG 973
QY 907 CAGCGGATGATCTTGTATCTGACGCTGATTAAGAAACAATCACTTTATGGGCTGAG 966
DB 974 CAGCGGATGATCTTGTATCTGACGCTGATTAAGAAACAATCACTTTATGGGCTGAG 1033
QY 967 AAGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1026
DB 1034 AAGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1093
QY 1027 CTCTTCATCTGCTATGAGAGCATGAGTGAACCAAGTGTACCCCTTATGATCCGATTG 1086
DB 1094 CTCTTCATCTGCTATGAGAGCATGAGTGAACCAAGTGTACCCCTTATGATCCGATTG 1153
QY 1087 TGACACTTTTGTGATGCTCATGAGAAATATGCTCAAGCTGAGAGAAATGAGAGCA 1146
DB 1154 TGACACTTTTGTGATGCTCATGAGAAATATGCTCAAGCTGAGAGAAATGAGAGCA 1213
QY 1147 CATTTGTTCCAACTTTGGAAGTATGTTCAAGTGTCAAGTGTCAAGTGTCAAGTGT 1206
DB 1214 CATTTGTTCCAACTTTGGAAGTATGTTCAAGTGTCAAGTGTCAAGTGTCAAGTGT 1273
QY 1207 AGTGGATGAGAGAGATCTTCAATAGAGATGAGAGAGATGAGAGATGAGAGAG 1266
DB 1274 AGTGGATGAGAGAGATCTTCAATAGAGATGAGAGAGATGAGAGATGAGAGAG 1333
QY 1267 GTGATATGAGAGATCTTCAATAGAGATGAGAGATGAGAGATGAGAGATGAGAGAG 1326
DB 1334 GTGATATGAGAGATCTTCAATAGAGATGAGAGATGAGAGATGAGAGATGAGAGAG 1393
QY 1327 GGAAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAG 1386
DB 1394 GGAAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAG 1453
QY 1387 TCTTGGGAGGAGGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAG 1446
DB 1454 TCTTGGGAGGAGGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAG 1513
QY 1447 CTTGGAATGCTCAGAGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGAGAG 1506
DB 1514 CTTGGAATGCTCAGAGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGAGAG 1573
QY 1507 TCTACTTGCACCAAGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGAGAG 1566
DB 1574 TCTACTTGCACCAAGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGAGAG 1633
QY 1567 GTGAGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAG 1626
DB 1634 GTGAGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAG 1693
QY 1627 CAGCACTTTTCTCAAGAGGCTTCTGTAACCAAGATGAGATGAGATGAGATGAGAG 1686
DB 1694 CAGCACTTTTCTCAAGAGGCTTCTGTAACCAAGATGAGATGAGATGAGATGAGAG 1753
QY 1687 CATCTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAGAG 1746
DB 1754 CATCTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAGAG 1813
QY 1747 GTGAGGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAGAG 1806
DB 1814 GTGAGGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAGAG 1873
QY 1807 TTTCTGATATATCCCAAAATTTCTTGAAGAGCTGAGTGTGAGAGATTTCTTTCAT 1866
DB 1874 TTTCTGATATATCCCAAAATTTCTTGAAGAGCTGAGTGTGAGAGATTTCTTTCAT 1933

Db 554 ATGGGTCGGATCGATTGGGCGGTGACACTTGTCTCTCACTCCGAAACCTGCAC 613
Qy 553 C---TTGGAGATTAGAGTTTGTCTGAAATACCTACCAAAATTCAGTGG---ATCAGAG 606
Db 614 CGACACAGAAATCTTGGGGCTGNTNGAACGTTTGCAGAGTTCAATTTGGGGTCAAGGAG 673
Qy 607 GAATGTTCTTGTATCTGATAGTGTGTGGAGTGTCCCTGTGCCATCACCCGATTGGA 666
Db 674 GGAATGTTCTCTCACTGATAGTGTGTGGAGTGTCCCTGTGCCATCACCCGATTGGA 733
Qy 667 GAATATGCTTGTGTGCTCGAGTACCTTATTTGTTCTTCTTGGTTGCAAAATTTT 726
Db 734 GGCTATGCTTGGAGTTCTGTGATCCCTCACTATNTACTTGGACTGATCTTCTTCT 793
Qy 727 ATCTTCCTGAATCTCCAGATGGCTGTAGCAAAAGTGGATGGAGAGCAAAAGAG 786
Db 794 ACTTGCCTGAATCAACCAAGGTGGCTGTNAGCAAAAGAGATGGCGAGGGCAAGAG 853
Qy 787 TGTTCGAAAGTTAAGGGGGAAGAGATGTCTCAGTGAATTTCTCTTCTTCTGAA 846
Db 854 TGAATGAAAGGCTCGGGGAGAGAGATGTCTCAGGAGAGAGGCTCTTCTTGAAG 913
Qy 847 GGTTCGAGTTGGAGGAGCACTTCCTGAAGAGTATCATTTGAGACCTGACCCAGAG 906
Db 914 GTTTGGGGGTGGTAAGATACAGTATTTNAGATCATTTGAGACCTGACCCAGAG 973
Qy 907 CAGCCGATATCTTGTACTGACGATATAGGAACAAATCACACTTTATGGGCTGAAG 966
Db 974 CAGCCGATATCTTGTACTGACGATATAGGAACAAATCACACTTTATGGGCTGAAG 1033
Qy 967 AAGCCAGTATGATTTGCTTCCGCTTTTCAAGGACCCATCATGCTTGAAGTGTCTT 1026
Db 1034 AAGCCAGTATGATTTGCTTCCGCTTTTCAAGGACCCATCATGCTTGAAGTGTCTT 1093
Qy 1027 CTCTTCATCTCGTCAATGGAGCATGTGTGAACAGAGTGTACCCCTTATGATCCGATT 1086
Db 1094 CTCTTCATCTCGTCAATGGAGCATGTGTGAACAGAGTGTACCCCTTATGATCCGATT 1153
Qy 1087 TGAACATTTTGTGTAGTGTCCATGAGATATGCTTAAGCTGAGAGATAGAGAGA 1146
Db 1154 TGAACATTTTGTGTAGTGTCCATGAGATATGCTTAAGCTGAGAGATAGAGAGA 1213
Qy 1147 CATTTGTTCCAACTTGTGAAGTATGTTCAGTGTCAAGATCAGCATGCCAAAATGAC 1206
Db 1214 CATTTGTTCCAACTTGTGAAGTATGTTCAGTGTCAAGATCAGCATGCCAAAATGAC 1273
Qy 1207 AGTGGAGTGAAGATCTTCATAGGGATGACAGAGATGACGATCTGATGTCAGAG 1266
Db 1274 AGTGGAGTGAAGATCTTCATAGGGATGACAGAGATGACGATCTGATGTCAGAG 1333
Qy 1267 GTGACTATGAGCAATCTTCATAGGCCATTTGCTGTCCAGGAGCAAGGTGCGAAG 1326
Db 1334 GTGACTATGAGCAATCTTCATAGGCCATTTGCTGTCCAGGAGCAAGGTGCGAAG 1393
Qy 1327 GGAAGACATTTGTGCACATGTCAACCGTGAAGTCTTTAGCATGAGAAGCAACC 1386
Db 1394 GGAAGACATTTGTGCACATGTCAACCGTGAAGTCTTTAGCATGAGAAGCAAGCC 1453
Qy 1387 TCTTAGGGAGGGTGAAGATGTGTAGAGCAGCTATATCGGTGGGGATGGCAGCTTG 1446
Db 1454 TCTTAGGGAGGGTGAAGATGTGTAGAGCAGCTATATCGGTGGGGATGGCAGCTTG 1513
Qy 1447 CTTCGAAATGTCAGAGAGAGAGTGAAGATGTGAAGAGAGAGGTGTTCCAAAGAG 1506
Db 1514 CTTCGAAATGTCAGAGAGAGAGTGAAGATGTGAAGAGAGAGGTGTTCCAAAGAG 1573
Qy 1507 TCTTACTTGAACCAAGGAGAGTCTTGGCTCAAGAGGGGCTCAATTTGTTCACTTCCG 1566
Db 1574 TCTTACTTGAACCAAGGAGAGTCTTGGCTCAAGAGGGGCTCAATTTGTTCACTTCCG 1633
Qy 1567 GTGGTGGCATGTTTGTAGGGGTAGTGTGTGATGCTGTGCTTTAGTAAGCACT 1626

Db 1634 GTGGTGGCATGTTTGTAGGGGTAGTGTGTGATGATGCTGTGCTTTAGTACTACT 1693
Qy 1627 CAGCACTTTTCTCAAGAGGTCTTGTGTGAACCAAGCATGTCAAGTCTGCCATGTTCAAC 1686
Db 1694 CAGCACTTTTCTCAAGAGGTCTTGTGTGAACCAAGCATGTCAAGTCTGCCATGTTCAAC 1753
Qy 1687 CATCTGAGGTAGCTGCCAAAGTTTCAAGTGAAGATTTGTTGAACCTGAGAGAGG 1746
Db 1754 CATCTGAGGTAGCTGCCAAAGTTTCAAGTGAAGATTTGTTGAACCTGAGAGAGG 1813
Qy 1747 GTGCCCTGTAGTGGGTGTGGAAATTCAGATCTTCAACAGTTTGTGGAAATTAACGGTG 1806
Db 1814 GTGCCCTGTAGTGGGTGTGGAAATTCAGATCTTCAACAGTTTGTGGAAATTAACGGTG 1873
Qy 1807 TTTCTGATATACCCCAAAATCTTGAAGCAAGCTGTGTGGCAATTTCTTCCAAAT 1866
Db 1874 TTTCTGATATACCCCAAAATCTTGAAGCAAGCTGTGTGGCAATTTCTTCCAAAT 1933
Qy 1867 TTTGCTGAGCTGGGATCAGATCATCTGATATGATGTTCTCTCACTACCTTACTAATG 1926
Db 1934 TTTGCTGAGCTGGGATCAGATCATCTGATATGATGTTCTCTCACTACCTTACTAATG 1993
Qy 1927 TTTCTGATGAGCTTTGCAAGCTGTCTTATGATCTTTCGGAAGAGTTTTGTGTC 1986
Db 1994 TTTCTGATGAGCTTTGCAAGCTGTCTTATGATCTTTCGGAAGAGTTTTGTGTC 2053
Qy 1987 TAGGCAATTCGAATCTTGAATGATCTTATGATATCTGTGTGTGTCATCTAATG 2046
Db 2054 TAGGCAATTCGAATCTTGAATGATCTTATGATATCTGTGTGTGTCATCTAATG 2113
Qy 2047 ATTTGGATACCTACCCAGTGTGCTCTCACCGTGAAGTATGCTGATCTTCT 2106
Db 2114 ATTTGGATACCTACCCAGTGTGCTCTCACCGTGAAGTATGCTGATCTTCT 2173
Qy 2107 GCTTCGTTATGGGATTTGTCATCCCAACATTTATGTCAGAGATCTTTCACCA 2166
Db 2174 GCTTCGTTATGGGATTTGTCATCCCAACATTTATGTCAGAGATCTTTCACCA 2233
Qy 2167 GGGTTCGTGGCTCTGATATGCAATTTGCTTTTACATTTGATGAGATATCATG 2226
Db 2234 GGGTTCGTGGCTCTGATATGCAATTTGCTTTTACATTTGATGAGATATCATG 2293
Qy 2227 TCACCTAAGCCCTTCGTGATGTCGATATGATGATGATGATGATGATGATGATGAT 2286
Db 2294 TCACCTAAGCCCTTCGTGATGTCGATATGATGATGATGATGATGATGATGATGAT 2353
Qy 2287 ATGCACTGATGCTTGAATTTCTTGTGTGTGCTTCTTAAAGTCCCTGAGCAAG 2346
Db 2354 ATGCACTGATGCTTGAATTTCTTGTGTGTGCTTCTTAAAGTCCCTGAGCAAG 2413
Qy 2347 GGAATCCCTTGAAGTTATACCGAATCTTGTGAGTGTGCGAAGCAAGCGCTGCA 2406
Db 2414 GGAATCCCTTGAAGTTATACCGAATCTTGTGAGTGTGCGAAGCAAGCGCTGCA 2473
Qy 2407 AAGCTTAATTTCTTGTGATCTTGTGTGCAATTTGCACTGATAGTAACTTGA 2466
Db 2474 AAGCTTAATTTCTTGTGATCTTGTGTGCAATTTGCACTGATAGTAACTTGA 2533
Qy 2467 GGGGTTTCAACCAAGAGCTGGAGATTAATTTGATTTGTGTAATGTAAGGAAGCA 2526
Db 2534 GGGGTTTCAACCAAGAGCTGGAGATTAATTTGATTTGTGTAATGTAAGGAAGCA 2593
Qy 2527 ACATCTGCTATGCTCTCAACCGGTAAAGAGTCCCTCAATGCAAAATGAGTGTG 2586
Db 2594 ACATCTGCTATGCTCTCAACCGGTAAAGAGTCCCTCAATGCAAAATGAGTGTG 2653
Qy 2587 TAAATTTGATATGATATTAATTAATTTGATATTTGATATTTGATATTTGATATTTG 2646
Db 2654 TAAATTTGATATGATATTAATTAATTTGATATTTGATATTTGATATTTGATATTTG 2713
Qy 2647 TATTCAGCTGATGTTGTGATGAATCTTGAACCAAGATGATATGATCTGATCTGA 2706
Db 2714 TATTCAGCTGATGTTGTGATGAATCTTGAACCAAGATGATATGATCTGATCTGA 2773

QY 2707 TGTATTAATATTCATATCTCAATTAAGAATAATATGTTTCTCAAAAAA 2757
 |||||
 DB 2774 TGTATTAATATTCATATCTCAATTAAGAATAATATGTTTCTCAAAAAA 2824

RESULT 3
 ABK51965
 ID ABK51965 standard; cDNA; 2601 BP.
 XX

AC ABK51965;

DT 27-AUG-2002 (first entry)

XX Soybean contig encoding A. thaliana-like sugar transport protein.

XX Soybean; Arabidopsis thaliana-like sugar transport protein;
 KM carbohydrate transport; grain filling; annual field crop;
 KW plant; gene; ss.
 XX
 OS Glycine max.

XX

EH Key Location/Qualifiers

FT CDS 175..238

FT /tag= a
 FT /product= "Soybean Arabidopsis thaliana-like sugar
 transport protein"

XX US6383776-B1.

XX 07-MAY-2002.

XX 14-APR-1999; 99US-0291922.

XX 24-APR-1998; 98US-083044P.

XX (DUPO) DU PONT DE NEMOURS & CO E. I.

XX Allen SM, Hiltz WD, Kinney AJ, Tinney SV;

XX WPI; 2002-451386/48.

XX P-PSDB; AAU97204.

XX New nucleic acid encoding plant sugar-transport proteins, useful for
 PT preparing transgenic plants with altered carbohydrate distribution -

XX Claim 3; Column 35-38; 54pp; English.

XX The present invention relates to the isolation of plant polynucleotide
 CC sequences encoding an Arabidopsis thaliana-like sugar transport
 CC protein or Beta vulgaris-like sugar transport protein. The
 CC polynucleotide sequences are useful for altering the level of sugar
 CC transport protein in plants, i.e. for control of carbohydrate transport
 CC and distribution in plant cells, e.g. during grain filling of annual
 CC field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
 CC carbohydrate flows and sugar transport. The polynucleotide
 CC sequences can also be used to isolate cDNA sequences and genes that
 CC encode homologues of the new proteins. The present sequence
 CC representing a contig assembled from various soybean cDNA clones
 CC encodes an Arabidopsis thaliana-like sugar transport protein.
 XX

XX Sequence 2601 BP; 648 A; 543 C; 644 G; 766 T; 0 other;

XX Query Match 33.9%; Score 940.2; DB 24; Length 2601;

XX Best Local Similarity 66.0%; Pred. No. 1.6e-212;

XX Matches 1476; Conservative 0; Mismatches 713; Indels 48; Gaps 6;

QY 177 ATGTGGGGGGTCTTGTGCGCATAGTGGCTCATGGGCAATCTATTGCGGGGTGG 236
 |||||
 DB 175 ATGAAGGTGGCGCTCTGTGCTATTGGCGCTTCATTGGTAATTCCTCCAAAGATGG 234
 |||||
 QY 237 GACAAATGCGACATCGCAGCTGCTGTCTGTATATAAGAAAGAAATTTCAATTGCAAAAT 296
 |||||

DB 235 GATAATGCTACCATCGCGGGCTAATGTTACATTAAGAAAGACCTTGCTTTGGAA-- 292
 QY 297 GAGCCCACTGGGAGAGGACTAATTGTGTCATGTCATTAATGGGCGCACCATGTGCT 356
 |||||
 DB 293 ---CAACTATGAAAGGCTGTGTGGGCAATGCTCCGATGTGGAGCAACGGTAATCAC 348
 QY 357 ACATTTCCGGGCCATTATACAGACTGATGGCCGACGCCCTATGCTTAATTCCTCTCA 416
 |||||
 DB 349 ACATGCTCTGGTCCATATAGCGGATTTGGCTCGGTGGGACCAACCATATATATCTCAT 408
 QY 417 ATTCTGACTTCTTCAAGCGGCTCATCATGCTATGCTCTTAATGCTATGCTGCTG 476
 |||||
 DB 409 GTGCTCTAATTTCTTGGGTGGTGGTGTGATGCTGTGGTCCCCAAATGTATGTGTGG 468
 QY 477 TTGGACGCTTGATGATGATTTGGTATTTGGCTTGCTGTCACGCTTGCTTTGATC 536
 |||||
 DB 469 TTGGGAGGCTACTTGAATGATTTGGGATTTGGCTTGCTGATCTTGTCCGGGTCTAT 528
 QY 537 ATTTGAGAAATAGCCCTTCCGAGATTAGAGGTTTGTGAATACACTACCAATTCAGT 596
 |||||
 DB 529 ATATCTGAAGCGCGCGCTGGAATTAAGGGGCTGTGAATACGCTTCTCAGTTCACT 588
 QY 597 GGATCAGAGAGATGTTCTTGTCAATGTCATGCTGTGTTGGAGATGCTCCTGTCCATCA 656
 |||||
 DB 589 GGCTCTGAGAGATGTTTGTGCTATGATGTTTGGCATGTCATTTAGTCCCGCG 648
 QY 657 CCCGATGGAATATATGCTGTGTGCTGCGGATGACCTTCAATTTCTTCTTGGTTTG 716
 |||||
 DB 649 CTTAGCTGAGGCTATGCTGGGGTCTGTATCTCTCTCTGATTTTGGATTTG 708
 QY 717 ACAATATTTTATCTTCTGAAATCTCCAGATGCTGTTAGCAAGATCGATGGCAGAG 776
 |||||
 DB 709 ACCATTTTCTTCTCCGAGTCTCTCGTGCTGTGTCAGCAAGAGAAAGATGCTCAG 768
 QY 777 GCAAAAAGGTGTGCAAAAGTTACGGGAGAAAGCATGCTCAGTGAATGTGCTCT 836
 |||||
 DB 769 GCTAAGAGGTGCTCAAAAGATGCGGAGAGAGATGTGTCAAGCATGTCATTTG 828
 QY 837 CTTCTGAAAGGTTGGAGTTGGAGAGACCTTCCATTGAAGATATACATTTGAGACCT 896
 |||||
 DB 829 CTGTTGAAGTCTGGGATGGGGGTATATATCTATGAAAGATATCAATTTGGCCT 888
 QY 897 GCCACCGAGGCGGATGATCTTGTACTGACGGTGAATTAAGAACAAATCACACTTAT 956
 |||||
 DB 889 GCTGACGATGTGGCTGATGTATCAATGACAGAGAAAGATTAATTTGATTAAT 948
 QY 957 GGGCTGAAAGAGCGCAGTCAATGATGCTGACCTCTTAAGGAGCCCATATGCTTGA 1016
 |||||
 DB 949 GGATCCCAAGCAGGCTTTCTTGTATCAAAACCTGTCACTGACAGAGTTCTAATGG- 1007
 QY 1017 AGTGTCTTCTCTTGCATCTGTCATGTGAGAGCATGTGAAACCAAGTGTACCCCTTATG 1076
 |||||
 DB 1008 -----CCTTGCCTCACACCATGAGCATCATCAACAAAGCATGCCCCCTCAG 1056
 QY 1077 GATCCGATTTGACACTTTTGTGTGTGTCATGAGAAATATGCTCA-----AGTGA 1130
 |||||
 DB 1057 GATCTCTGTGACACTGTTTGTGACATTCATGAGAACTCCCGACAGAGAGACAGA 1116
 QY 1131 GGAAGTATGAGAGACATTTGTTCCAACTTTGGAAGTATGTCAGATGTCAGATGAG 1190
 |||||
 DB 1117 GGAAGCATGCGAAGCACTGTGTTCCAAATTTTGAAGCATTTGAGCATGCTGAGCGG 1176
 QY 1191 CATGCCAAATATGACAGTGGATGATGAAGAAATCTTCATAGGATGACAGAGATGACGA 1250
 |||||
 DB 1177 CATGCTAAATTTGAACAATGGATGAAGAACTTCAAGAGGAACGTGAGACATACAG 1236
 QY 1251 TCTGATGTCAGAGAGTGAATGATGAGGCAATCTCCATAGCCCATTTGCTCCAGGCG 1310
 |||||
 DB 1237 TGAAGTGAACCCGTGGGACCTCCGATGATTAATTTGACAGAGTCTTTAATCTCAGCC 1296
 QY 1311 GCACAGGTGCGAAGAGGAGACATTTGTGACCATGTGTCAACGTGGAAGTGTGAGC 1370
 |||||
 DB 1297 ACAACAGGCTTGAAGAAAGATTAACCTCTCTCTCCATGGAGATATCTTGGCAGC 1356

QY 1371 ATGAGAAGCA---AACCTTTAGGGAGGGTGGATGCTGTGACGACACTGATATC 1427
 DB 1357 ATGAGCGGTACAGATGATGCTCATGCAAGGGTCAAGGGTACCAAGGGTGTGTACAGGATAT 1416
 QY 1428 GTTGGGGAGTGGCAGCTTCTTGAAATGATGACAGAAAGAGTGAATGATGAAG 1487
 DB 1417 GGTGGTGGCTGGCACTGGCACTGGAATGGAATGACTGATA---AGTGTAGATGGAACCA 1473
 QY 1488 GAAAGTGTTCACAAAGAGCTTACTTGACCAAGAGGAGTTCCTGGCTCAAGAAAGGGC 1547
 DB 1474 CAAGAGGGGTTAAAGATTTATTTACATAGAGAGGAGTTCCTGCTGCTGGA 1533
 QY 1548 TCAATTTTCACTCCCGGTGTGGCCATGTTTTGAGGGATGAGATTTGTACATGCT 1607
 DB 1534 TCCATTGATTCATTCCTCCGTGAAGGCGAA-----TTTGTCCAGGCT 1575
 QY 1608 GCTGCTTTAGTAACTCAGACACTTTCTCAAAAGGCTTGTGTAACACGACATGTCA 1667
 DB 1576 GCTGCTTTAGTAACTCAGACACTTTCTCAAAAGGCTTGTGTAACACGACATGTCA 1635
 QY 1668 GATGCTGCCATGTTCAACCATCTGAGTAGCTGCCAAAGGTTCAAGTTGAAGATTTG 1727
 DB 1636 GGGCTGCAATGCTTCAACCATCTGAGTAGCTGCCAAAGGTTCAAGTTGAAGATTTG 1695
 QY 1728 TTTGAACCTGAGTAGGGGTGCTGTAGTGGGTGTTGAATTCAGATCTTCAACG 1787
 DB 1696 CTTGACCAAGGGGTGAGTATGATGATGATGATGATGATGATGATGATGATGATG 1755
 QY 1788 TTTGCTGGAATAAGGTTGCTGTACTATACCCCAAAATCTTGAGCAAGCTGATG 1847
 DB 1756 TTTTACAGGATTAATGGGTTTATATATACACCCCAAAATCTTGAGCAAGCTGATG 1815
 QY 1848 GCAATTTTCTTCCAAATTTGCTGCTGAGCTGGCCTGAGCATTCATCTTATCAGTT 1907
 DB 1816 GAAATTTTCTTCCAAATTTGCTGCTGAGCTGGCCTGAGCATTCATCTTATCAGTT 1875
 QY 1908 CTGACTACCTTACTATGCTTCTGCTGAGCTGGCCTGAGCATTCATCTTATCAGTT 1967
 DB 1876 TTTGACCAAGGGGTGAGTATGATGATGATGATGATGATGATGATGATGATGATG 1935
 QY 1968 GGAAGAGGTTTCTGCTGCTGAGCTGGCCTGAGCATTCATCTTATCAGTT 2027
 DB 1936 GGAAGAGGTTTCTGCTGCTGAGCTGGCCTGAGCATTCATCTTATCAGTT 1995
 QY 2028 GTTGTGCTCAATCTATGATTTGGGTACACTAGCCCATGCTTGTCTTCAACCGTCA 2087
 DB 1996 GTTGTGCTCAATCTATGATTTGGGTACACTAGCCCATGCTTGTCTTCAACCGTCA 2055
 QY 2088 GTTGTGCTCAATCTATGATTTGGGTACACTAGCCCATGCTTGTCTTCAACCGTCA 2147
 DB 2056 GTTGTGCTCAATCTATGATTTGGGTACACTAGCCCATGCTTGTCTTCAACCGTCA 2115
 QY 2148 GCAGAGATCTTTCACCAAGGGTGTGGCTCTGATATGAGCAATTTGAGCTTATCATT 2207
 DB 2116 TCAAGAGATTTTCCCACTAGGGTGTGGCTCTGATATGAGCAATTTGAGCTTATCATT 2175
 QY 2208 TGGATCGAGATATCATGCTGCTGCTGAGCTGGCCTGAGCATTCATCTTATCAGTT 2267
 DB 2176 TGGATCGAGATATCATGCTGCTGCTGAGCTGGCCTGAGCATTCATCTTATCAGTT 2235
 QY 2268 GCGGCTGTTTTCAGATATGAGCTGATGCTGATGCTTGTGTGTGTCTTCTTCT 2327
 DB 2236 GCGGCTGTTTTCAGATATGAGCTGATGCTGATGCTTGTGTGTGTCTTCTTCTTCT 2295
 QY 2328 AAGGCTCTGAGACAAAGGGATGCGCTTGAAGTTATTAACGAATCTTTGAGATGGT 2387
 DB 2296 AAGGCTCTGAGACAAAGGGATGCGCTTGAAGTTATTAACGAATCTTTGAGATGGT 2355
 QY 2388 GCGAAGCAAGCGGCTGC 2404
 DB 2356 GCGAAGCAAGCGGCTGC 2372

RESULT 4
 ABX93201
 ID ABX93201 standard; cDNA, 2601 BP.
 XX
 AC ABX93201;
 XX
 DT 29-MAY-2003 (first entry)
 XX
 DE cDNA encoding soybean sugar transport protein #1.
 XX
 KW Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
 KW plant sugar transport protein; carbohydrate transport; soybean;
 KW carbohydrate distribution; plant; gene; ss.
 XX
 OS Glycine max.
 XX
 PN US2002178468-A1.
 XX
 PD 28-NOV-2002.
 XX
 PF 17-JAN-2002; 2002US-0051902.
 XX
 PR 24-APR-1998; 98US-083044P.
 PR 14-APR-1999; 99US-0291922.
 XX
 PA (ALLEN/) ALLEN S. M.
 PA (HITZ/) HITZ W. D.
 PA (KINNEY/) KINNEY A. J.
 PA (TINGEY/) TINGEY S. V.
 PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
 XX
 DR WPI, 2003-340957/32.
 DR P-PSDB; AB08329.
 XX
 PT Novel plant sugar transport proteins and nucleic acid encoding the
 PT protein useful for producing transgenic plants having altered levels of
 PT sugar transport protein
 XX
 PS Claim 2; Page 19-20; 56pp; English.
 CC The present invention relates to the isolation of Arabidopsis
 CC thaliana-like or Beta vulgaris-like sugar transport proteins, and
 CC the polynucleotide sequences encoding them. The plant sugar transport
 CC proteins of the invention have been isolated from corn, rice, soybean,
 CC and wheat. The polypeptides of the invention may be used for altering
 CC the level of expression of a sugar transport protein in a host cell,
 CC by transforming a host cell with a chimeric construct encoding all,
 CC or a portion of the sugar transport protein, in sense or antisense
 CC orientation. Particularly, the polypeptides may provide a means to
 CC control carbohydrate transport and distribution in plants.
 CC ABX93198-ABX93205 represent cDNA sequences encoding Arabidopsis
 CC thaliana-like sugar transport proteins.
 CC
 SQ Sequence 2601 BP; 648 A; 543 C; 644 G; 766 T; 0 other;
 Query Match 33.9%; Score 940.2; DB 25; Length 2601;
 Best Local Similarity 66.0%; Pred. No. 1.6e-212;
 Matches 1476; Conservative 0; Mismatches 713; Indels 48; Gaps 6;
 QY 177 ATGTCGGGGGGTGTCTTGTGTCATGATGCTGCTTCCATCGGCAATCTATTCAGAGGGTGG 236
 DB 175 ATGAAAGGTGCGCTGTGTGCTATGCTGCTTCCATGTGTAATTTCTCCAAAGAGG 234
 QY 237 GACATGCAACCATCGAGCTGCTGCTGTAATTAAGAAAGATTTCAATTCGAATAT 296
 DB 235 GATATGCTACATCGCGGGGCTAATGCTTAAGAAAGCTTGTGGAA-- 292
 QY 297 GAGCCACTGTGAGGAGCTAATTTGTCAATGCTCAATTCAGGCGCCACCATCTTACT 356
 DB 293 ----CAACTATGAAAGGCTTGTGTGGGCAATGCTCGATTTGAGCAACGATATACCC 348

QY 357 ACATTCCTCGGCGCATTTATCAGACTGATGCGGCGGACGCCCTATGCTTATCTCTTCA 416
DB 349 ACATGCTCTGATCTTATATAGCGATTTGGCTCGGTCGGCGACCATATGATATATCTATCT 408
QY 417 ATTCTGACTTTCTTCAAGCGGCTCATCATGCTATGCTCTTAACTCTATGCTCTG 476
DB 409 GGTCTCATATTTCTTGGGTGGTTGGTATGCTGTGCTCCCAAAATGATGATGTTGTC 468
QY 477 TTGGACGCTTGATGATGATTTGGTATTTGGCTTGGCTGCTGACCGCTTGCTTGTAC 536
DB 469 TTGGGAGGCTACTTGATGATTTGGGATTTGGCTGCTGCTGCTGCTTGTCCCGCTAT 528
QY 537 ATTGAGAATATGCCCCCTTCGAGATATGAGTTTGTCTGATATCACTACCAATTCAT 596
DB 529 ATATCTGAACGCGCCGCTGTAATTAAGGGGTGGTTGATATAGCTTCTCAGTTCACT 588
QY 557 GGATCAGAGGATGTTCTTGTCAATCTGCAATGCTGTTGGAGATGCTCTGCGCCATCA 656
DB 589 GGCTCTGAGGATGTTTGTCTGATGCTATGCTTGGCTTGGCAGTCAATGAGTCCCGC 648
QY 657 CCCGATGAGAAATATGCTTGGTGGCTCGGATACCTTCATGTTCTTGGTTTG 716
DB 649 CTTAGCTGAGGCTCATGCTTGGGTTCTGCTATCTCTCTTGTATTTGGATTG 708
QY 717 ACAATATTTTATCTTCTGATATCTCCAGATGCTCTGTTAGCAAAAGTGCATGCGAG 776
DB 709 ACCATTTTTTTCTTGGCCGAGTCTCTCGTGGCTGTGTCAGCAAGAGATGCTCGAG 768
QY 777 GCAAAAAGTGTGCAAAAAGTTACGGGGGAAAGAGATGCTCAAGTGAATTTGCTCT 836
DB 769 GCTAAGAGGTGCTCCAAAGATTTGGCGGAGAGGATGCTCAGGCGAGATGCGATTG 828
QY 837 CTTCTGAAAGGTTGAGGTTGAGGTTGAGGACACTTCATTTGAAGATGATCAATGGA 896
DB 829 CTGTTGAAGTCTCGGATTTGGGGGTGATCATCTATGAAAGATCAATTTGGCCT 888
QY 897 GCCACGAGGACCGGATGATCTTGTACTGACGGTATGAGAAATCAATCACTTAT 956
DB 889 GCTGACGATGCTGATGCTATGATGAACATGCAACGAGAAATTAATTTGATTAAT 948
QY 957 GGGCTGAAGAGCCAGTCAATGATGCTGACCTTTGAAGGACCCATCATGCTTGA 1016
DB 949 GGATCCCAAGCAGGCTTCTGTTATCAAAACCTGTCAGACAGAGTTCTAATGG- 1007
QY 1017 AGTGTCTTCTCTGCACTCTGTGATGAGGACATGCTGAACCAAGTGTACCTTATG 1076
DB 1008 -----CTTGGCTGACACCATGAGATCATCAACCAAGCATGCCCCCTCATG 1056
QY 1077 GATCCGATTTGACACTTTTGGTATGTCATGAGATATGCTCA-----AGCTGA 1130
DB 1057 GATCTCTGTGACACTGTTTGGTATGATTTCAATGAGACCTCCCGAGACAGAGCAAG 1116
QY 1131 GGAAGTATGAGAGCAATTTGTTCCAACTTTGGAAGTATTTCAAGTGTCAAGATCAG 1190
DB 1117 GGAAACATGCGAAGCACTGTGTTCCAAATTTTGAAGCATGTTGACACTGCTGAGCG 1176
QY 1191 CATGCAAAAATGACAGATGAGATGAAGATCTTCAATGAGGATGACAGAGATGACGA 1250
DB 1177 CATGCTAAATTTGAACAATGAGATGAAGAACTTACAAAGGGAACGTATGAGATCAATG 1236
QY 1251 TCTGATGTCAGAGAGTACTATGAGACATCTCCATAGCCCATTTGCTGTCCAGGCG 1310
DB 1237 TCAAGTGAACCCGTGGGACCTCCGATGATTAATTTGACAGTCTTTAATCTCAGGCCA 1296
QY 1311 GCAACAGGTGCGAAGGAGAGACATTTGACCAATGATCAACGTGAGAGTGTGAGC 1370
DB 1297 ACAACAGGCTTGAAAAMAAGATTACCTCTCTCTCTTCCATGGCAGTATCTTGGCAGC 1356
QY 1371 ATGAGAGGCA---AACCTTTAGGAGAGGTGAGATGCTGAGCAGCACTGATATC 1427
DB 1357 ATGAGGCTGACAGTATGCTCATGCAAGGCTCAGGTCAGCAAGGTGTGTACAGGTAT 1416
QY 1428 GGTGGGAGTGGCAGCTTCTTGAAATGCTCAGAGAAAGAAAGTGTAGAAAG 1487

DB 1417 GGTGTGGCTGCAACTGGCATGGAATGACATGATA---AGTGAAGATGGAACAA 1473
QY 1488 GAAAGTGTTCAAAAGAGTCTACTTGCACCAAGAGGAGTTCTGCTCAAGAGGAGC 1547
DB 1474 CAAGAGGAGTTTAAAGATTTATTTACATGAGGAGGAGTTTCTGCAATCTGTGTGGA 1533
QY 1548 TCAATTTTCACTCCCGGTGGTGGCGATGTTTGAAGGATGAGATTTGATATGCT 1607
DB 1534 TCCATTTATGATTCCTCCGATGAGGCA-----TTTGTCAAGCT 1575
QY 1608 GCTGCTTATGATGACAGTCTGACACTTTTCTCAAGGCTCTTGTGAACCAAGCATGTCA 1667
DB 1576 GCTGCTTGTATGAGCAACCGCTCTTTACTCAAGAGCTTATGATGACACCAAGT 1635
QY 1668 GATGCTGCATGTTTCACTCATCTGAGGTAGTGCACAAAGTTCACTTGAAGATTTG 1727
DB 1636 GGGCTGCAATGTTTCACTGAGCAGCTTCAAGAGGAGGCAAGTTGGAAGCTCTT 1695
QY 1728 TTTGAACCTGAGATGAGGCGGCTGTTAGTCGCTGTGGAATTCAGATCTTCAAG 1787
DB 1696 CTTGAACAGGAGTTAAGCATGATGATGTTGAGTTGGAATTCAGATCTTCAAG 1755
QY 1788 TTTGCTGAAATTAACGAGTGTCTGTACTATACCCCAAAATTTCTGAGCAAGCTGTG 1847
DB 1756 TTTTCAAGGATTAATGGGTTCTATATTAACACCTCAATCTTGAAGAGCGGTGT 1815
QY 1848 GCAATTTCTTCCAAATTTGCTCAGCTGCGCATGACATTCATCTTATGATGCT 1907
DB 1816 GAAATTTCTTCCAGATTAATGAGCAATTTGCTCAGAGTGGCATTCATCTTATCAGTCT 1875
QY 1908 CTCACTACTTACTATGCTTCTGCTGATGCTTGGCTTGGCATGCTGATGATGCTTCC 1967
DB 1876 TTTCAACCTTCTGATGCTTCTGATTAATGAGGCTGAGGAGCATGATGATGATTTCA 1935
QY 1968 GGAAGAGTTTTCTGCTGAGGACAAATTCATCTTATGATGATCTTATGCTG 2027
DB 1936 GGCAGAGGAGTGTCTACTTACTTACTTCAATCCCGTGTGATGCTACTATTTTGG 1995
QY 2028 GTTGTGTCATCTAATTTGATTTGGGTAACATGAGCCATGCTTGTCTCAACGCTCAGT 2087
DB 1996 GTCATTTGAGGCTGTGTAATTTTGGCAATGAGGCTGAGCAATCTCAACAGTATGC 2055
QY 2088 GTTATGCTACTTCTGCTGCTGCTGATGAGATTTGGTGTCCATGCCAATTTTATGT 2147
DB 2056 GTTGTGTTTATTTCTGCTGCTTGTGTGATGAGGTTATGACCAATTCCAAACTTTTGC 2115
QY 2148 GCAGATCTTTCACACAGGAGTTCGAGGCTCTGATATTTGCAATTTGACCTTATCATTT 2207
DB 2116 TCAAGATTTTCCCACTAGGAGTGTGCTGCTGATGCTGATGCTATGATGTTC 2175
QY 2208 TGGATGAGATATCATGCTCACTGACCTTCTGTGATGCTGAATGATTTGAGCTG 2267
DB 2176 TGGATTTGAGACATCATCATCACTGCTGCTGTGATGCTGAGCTTTTGAAGACTT 2235
QY 2268 GCGGTGTTTTCAGATATATGACGCTGATGCTTGAATTTCTTGTGTGTCTCTT 2327
DB 2236 GGTGTGATTTGCTTATGAGCAGTTTGTGTTCACTGATGATTTGTGTTTGG 2295
QY 2328 AAGTCCCTGAGACAAAGGAGTGCCTTGAAGTTATTAAGGATTTCTTGAAGTGT 2387
DB 2236 AAGTTCCAGAAACAAAGGAGTGCCTTGAAGTATCTGATTTCTTTCTGTGGA 2355
QY 2388 GCGAAGCAGCGCTGC 2404
DB 2356 GCAAGCAGGCTGCTTC 2372
RESULT 5
AB213510
ID AB213510 standard; DNA; 2190 BP.
XX
AC AB213510;

XX	21-JAN-2003	(first entry)
DT	Arabidopsis thaliana stress regulated gene SEQ ID NO 1315.	
XX		
DE	Arabidopsis thaliana, plant; gene; stress; transgenic; ds.	
XX		
KW	Arabidopsis thaliana.	
XX		
OS	Arabidopsis thaliana.	
XX		
PN	WO200216655-A2.	
XX		
PD	28-FEB-2002.	
XX		
PF	24-AUG-2001; 2001MO-US266685.	
XX		
PR	24-AUG-2000; 2000US-227866P.	
XX		
PR	26-JAN-2001; 2001US-264647P.	
XX		
PR	22-JUN-2001; 2001US-300111P.	
XX		
PA	(SCRI) SCRIPPS RES INST.	
XX	(SYGN) SYNGENTA PARTICIPATIONS AG.	
XX		
PI	Harper JF, Kreps J, Wang X, Zhu T;	
XX		
DR	WPI; 2002-304127/34.	
XX		
PT	Identifying a stress condition to which a plant cell has been exposed	
XX	and producing plants with increased tolerance to these abiotic stresses	
PT	-	
XX		
PS	Claim 144; SEQ ID NO 1315; 577pp + Sequence listing; English.	

CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising:
CC (a) contacting nucleic acid representative of expressed polynucleotides
CC in the plant cell with an array or probes representative of the plant
CC cell genome; and
CC (b) detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
CC in methods of the invention.
CC Note: The sequence data for this patent is not represented in the printed
CC specification but is based on sequence information supplied to Derwent by
CC the European patent Office.

Sequence 2190 BP; 529 A; 439 C; 572 G; 650 T; 0 other;

Query Match	31.5%	Score 875.8;	DB 24;	Length 2190;
Best Local Similarity	64.9%	Pred. No. 2.8e-197;		
Matches 1449; Conservative	0;	Mismatches 722;	Indels 60;	Gaps 8

OY	177	ATGTGGGGGGCTGTCTTGTGCGCAGATGAGGCTCCATGGCAATCTATATGCAAGGGGTGG	236
Db	1	ATGAGTGGAGCTGTGCTTTGTCTATATGCTGCTGCTGTGGCACTTTTACAAGATGG	60
OY	237	GACAAATGCCAACATGCACTGCTGTGTTCTGTATATAAAGAGAAATTTCAATTGCCAAAT	296
Db	61	GATTAACGCAACTATTCACAGAGCTGTGTTGTATATAAAGAGTTAATTTGGAGAGT	120
OY	297	GAGCCACTGTGGAGGGACTAATTTGTCTAATGTCAATTATGGCGGCACATTCGTACT	356
Db	121	AATTCATCACTGAGTGAAGGTCTAATTTGTGGCCATGTCAATTTGTGTACTCTGATTTACA	180
OY	357	ACATTTCTCGGGCCATTATCAGACTCGATTGGCGGAGGCCCTTATTCCTCTCTCA	416
Db	181	ACATGCTCTGGAGGGGTAGCTGATTTGGCTTTGGTCGCCGTCCCATATGTAATTTGCTCA	240
OY	417	ATTCTGACTTCTTACAGGGGCTCATCATGCTATGTTCTCTAATGTCTATGTCTGCTG	476
Db	241	ATTCTCTACTTGTGTGGTTCTCTAGTAATCTATGTGCTCCGATATGTTATGTGTGCTC	300

QY	477	TTGGCAGCGCTCGAGATGGAATTGGATATGGCGTTGGCTGTCAGCGTTGGCCCTTTGAC	536
Db	301	TTAGGAAGTGTTAGAGATGATTTGGGGGTGGTCTTGCGTCACACTGTGTCTCAATTTAT	360
QY	537	ATTTCAGAAATAGCGCCCTTCGGAGATTAGAGGTTTGGCTGAATACACTACCAATTCAGT	596
Db	361	ATTATCTGAGACTGCACCACCTTGABATTAGGGGACTGTTGAAATAGCTACCCGACTTCACT	420
QY	597	GGATTCAGAGGAATGTTCTTGTCATPACTGCATGCGTGTGGATGTGCTGTGGCCATCA	656
Db	421	GGCTCTGGAGGGATGTTCTTATCTTACTGTATGTTTTGGAAATGTCGTTGATGCATCA	480
QY	657	CCCGATTGGAATATTACCTTGGTGTGCTGGCAGTACCTTCATGTCTTCTTGGTTTG	716
Db	481	CTTAGCTGGAATTAATGATGCTTGGTGTCTTTTCACTCCCTTCTTCTTCTTCTCTC	540
QY	717	ACAATATTATTATCTTCTCGAATCTTCCAAGATGGCTCGTTAGCAAGGTCGATGGCAGAG	776
Db	541	ACGGTCTTCTTCTTGCCCGAGTCCCCAAGTGGCTCTGTGACCAAAAGTGCATTCCTTGA	600
QY	777	GCAAAAAAGTGTTGCAAAAGTTACGGGGGAAAGACGATGTTCTCAGGTGAATGTCCCTT	836
Db	601	GCAAAAGCGGGTCTTTCAGAGACTCGTGTGTCGCAAGATGTTCTGGGAGATGGCTTTG	660
QY	837	CTTCTCCAGAGGCTTGAGGTTGGAGAGACACTTCCATTGAAGTACATCATTTGAGACT	896
Db	661	TTGGTGTGGGCTCTTGGAATTTGAGGGTGAAACCAATAGGAATATATAATTTGGTCCC	720
QY	897	GCCACCGAGGCGAGCGAGTATCTTGTAACAGCGTGTATAGAAACAAATCACACTTAT	956
Db	721	GCGATGGAATTACTGATGATCATGATATAGCTGTGGATTAAGATTCAAATTAAGTTATAT	780
QY	957	GAGGCTGGAAGAGGCCAGTCAATGATGTCCTGCACTTTCAAGGAGCCCATCATGCTTTGA	1016
Db	781	GGTGAGAGAAAGGGCTGAGTTGGGTTGCTTAGGCCAATCAAAAG-----GGA	828
QY	1017	AGTGTGCTTCTCTTGCACTCTGTCATGGAAGCATGTTGAACCAAG-----TGTAACCTT	1073
Db	829	AGCATATAGATGTTTGTCTCGCCATGGAAGTCAATAGACAGAGGCAAGGCTCATTG	888
QY	1074	ATGATCCGATTGTCACCTTTTGGTAGTGTCATAGAAATATGCCCAAGCTGAGGA	1133
Db	889	ATTGATCTCTTGTCACCTGTGTTGGAGCTTTCAGGAAGATGCGGACACT---GGA	945
QY	1134	AGTATGAGGACACATGTTTCCAAACTTTGGAAGTATGTTCAAGTGTACACATCAGAT	1193
Db	946	AGCATGAGGAATGCTTGTCTTCCACATTTTGGAGTATGTTCAAGTGTGAGGGAAATCA	1005
QY	1194	GCCAAAAATGAGCAGTGGAGTAAAGAAATCTTCAATAGAGATGACGAGAGTACGACT	1253
Db	1006	CCAAGACATGAAGATTGGAGTAGAAGAAATCTTGTGGAAGAAAGGTAGATTAATCACTC	1065
QY	1254	GATGTGAGAGAGTGAATATAGAGACAAATCTCATAGCCATTGCTGTCCAGCGAGCA	1313
Db	1066	GACCAT---GGAAGATGTTGAAGATGATTTCAATCTCCGTGATGTCAAGTCAAAAG	1122
QY	1314	ACAGGTGCGAAGGGAAGACATTGTGCAACCATGGTCAACCTGGAAGTCTTTGAGCATG	1373
Db	1123	ACAAGCATGGA---GAAAGCATGCTCTCACCTGCTCATGGAATCTTTTCTACCTTCGA	1179
QY	1374	AGAAAGCAAACTCTTAGGGGAGGGTGAAGATGCTGTAGACAGACCTGATATCGGTGG	1433
Db	1180	CATGAATCAAGTGCAGAGGAGCTCAAGGGGAAAGAGCGGGTAGATGGGAGATTGAGGT	1233
QY	1434	GGATGGCAGCTTGTCTTGAAATGTCAGAGAAAGAAAGGTGAGATGTATGAAAGGAAGT	1499
Db	1240	GGATGGCAAGTGGCATGGAATGACGGAAGAGAAAGATGATCGGGAAGAAAGAAAGAA	1299
QY	1494	GTTTTCAAAAGAGTCTACTTGCACCAAGAGGGAAGTCTGTGCTCAAGAAAGGGCTCAAT	1553
Db	1300	GGTTTC-----CCAGATCTCGACGTGGCTCAATTT	132
QY	1554	GTTTCACTCCCGGTGTGGCGATGTTTTTGGAGGTAGTGAATTTGTATCATGCTGCTGCT	1613

```

Db      1330 GTTTCATTGCTGGTGGTATGGAACCGGTGAGCA---GATTGTGACAAAGCTTCTCT 1386
Qy      1614 TTAGTAAGTCAATGCACTTTTCTCAAGGCTTCTGTCGAAACCGCATGTCAGATGCT 1673
Db      1387 TTGGTTAGCAACCAAGCTCTTATTCAAAGACCTTCTCAAGAAATATCAAAATGGTCTCT 1446
Qy      1674 GCCATGTTCAACCACTGAGGAGTGCACAAAGTTTCAAGTTGAAAGATTGTTTGA 1733
Db      1447 GGTATGTATCATCATCCGAAACAAT---AAAGGTCAATTTGGATATCTTCATGAT 1503
Qy      1734 CCTGAGTGAAGGCGTCCCTGTTAGTGGTGTGGAATTCAGATCTTCAACAGTTTGT 1793
Db      1504 CCTGAGTGAAGGCGTCCCTGTTAGTGGTGTGGAATTCAGATCTTCAACAGTTTGT 1563
Qy      1794 GGAATTAAGGCTGTTCTGTATCAATACCCCAAAATTTGAGCAAGCTGTGTGAGAT 1853
Db      1564 GGCATCAAGGAGTCTTACTACACACCGCAATCTTGTAGAGCGGCTGTGCGGATC 1623
Qy      1854 ATTTCTTCAAAATTTGCTCAGCTGGGATCAGCATCATCTTATGATCAGTTCTCTACT 1913
Db      1624 CTACTATGACAAATGGGATGATTTCTTCTCAGCATCTTACTTATTAAGTCAATGACA 1683
Qy      1914 ACCTTACTAATGCTTCTTGCATTTGCTTGGCCATGCTGCTTATGAGATCTTCCGGAGA 1973
Db      1684 ACCTTGTGATGTAACTGCAATAGCTGTGCAATGAGGCTCATGATCTTCTGTGCTGA 1743
Qy      1974 AGCTTTTCTGCTGAGGCAATTCGAATCTTATGATGATCTTCTGATTTCTGCTGTTG 2033
Db      1744 AGGACCTGCTTCTCAACGATACCAATCTGATGATCTTATGATTTGTTTATGATAC 1803
Qy      2034 TCCAAATCTAATGATTTGGGTAACATAAGCCCAATGCTTCTCCACGCTCAAGTGTATC 2093
Db      1804 TCAATCTTGTTCATGATGAACGATATGTCAGCGGCTTATCAACCGTAAAGCTTTG 1863
Qy      2094 GTCTACTTCTGCTGCTTCTGTTAAGGATTTGCTCCATCCCAATTTTATGTGACAG 2153
Db      1864 CTCTACTTCTGCTTCTGTTAAGGATTTGCTCCATCCCAATTTTATGTGACAG 1923
Qy      2154 ATCTTCTCAACCAAGGCTTGTGCGCTCTGTATGTCATTTGTCCTTTAATTTGATC 2213
Db      1924 ATTTTCTCAACCTGATCCGCGAATCTGATCCATGCGCATCTGCGCATCTTCTGATC 1983
Qy      2214 GGAGATATCATGCTGCTCAACAGCTTCTGTCGATGATGATGATGATGATGATGATG 2273
Db      1984 TGTGACATATGCTCACTTACAGCTCTCCCGTCTGCTCAATTCATGACATGATG 2043
Qy      2274 GTTTTCAGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2333
Db      2044 GTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2103
Qy      2334 CCTGAGACAAAGGAGTGCCTTGAAGTTATTAACGAATCTTTCAGATGTCGAGAG 2393
Db      2104 CCGGAATCTAAAGGAGTGCCTTGAAGTTATTAACGAATCTTTCAGATGTCGAGAG 2163
Qy      2394 CAAGCGGCTGC 2404
Db      2164 CAAGCTGAAGC 2174

```

```

RESULT 6
AB214449
ID      AB214449 standard; DNA; 2205 BP.
XX
AC      AB214449;
XX
DT      21-JAN-2003 (first entry)
XX
DE      Arabidopsis thaliana stress regulated gene SEQ ID NO 2254.
XX
KW      Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS      Arabidopsis thaliana.

```

```

XX      XX
PN      WO200216655-A2.
XX      28-FEB-2002.
PD      24-AUG-2001; 2001WO-US26685.
XX      24-AUG-2000; 2000US-227866P.
PR      26-JAN-2001; 2001US-264647P.
PR      22-JUN-2001; 2001US-300111P.
XX      (SCRI) SCRIPPS RES INST.
PA      (SYGN) SYNGENTA PARTICIPATIONS AG.
XX      Harper JF, Kreps J, Wang X, Zhu T;
PI      WPI; 2002-304127/34.
DR      Identifying a stress condition to which a plant cell has been exposed
PT      and producing plants with increased tolerance to these abiotic stresses
PT      -
XX      Claim 144; SEQ ID NO 2254; 577bp + Sequence Listing; English.
PS      The invention relates to identifying a stress condition to which a plant
XX      cell has been exposed, comprising:
XX      (a) contacting nucleic acid representative of expressed polynucleotides
XX      in the plant cell with an array of probes representative of the plant
XX      cell genome; and
XX      (b) detecting a profile of expressed polynucleotides in the plant cell
XX      characteristic of a stress response. The method is useful in the
XX      production of transgenic plants, cells and seeds and in producing plants
XX      with increased tolerance to abiotic stress. The present sequence is that
XX      of an Arabidopsis thaliana stress regulated gene (AB212156-AB217574) used
XX      in methods of the invention.
XX      CC Note: The sequence data for this patent is not represented in the printed
XX      specification but is based on sequence information supplied to Derwent by
XX      the European Patent Office.
SQ      Sequence 2205 BP; 521 A; 524 C; 521 G; 639 T; 0 other;
XX
Query Match      25.2%; Score 700; DB 24; Length 2205;
Best Local Similarity 61.2%; Pred. No. 1.2e-155;
Matches 1371; Conservative 0; Mismatches 770; Indels 99; Gaps 11;
Qy      177 ATGTCGGGGGCTGTTCTTGTGCGCATGATGCTGCTTCATGCGCATTTGAGGGGTGG 236
Db      1 ATGAAGGAGAGCACTCTGTTGCTCTCGCGCAATGCGCAATTTCTTACAAAGATGG 60
Qy      237 GACATATGCCACCATGCGAGCTGCTGCTGTATATAAAGAGATTTCAATTTGCAAAAT 296
Db      61 GACATATGCCACCATGCGAGCTGCTGCTGTATATAAAGAGATTTCAATTTGCAAAAT 118
Qy      297 GAGCCACTGTGAGAGGACTAATTTGTCAATATGCTATTAATGCGGCCACCATCTTACT 356
Db      119 -----CTCTGTTCAAGGCTTGTGCTGTATGATGATGATGATGATGATGATGATG 174
Qy      357 ACATTTCTCGGGGCAATTAACAGATCGATGATGCGGAGCGCTTATGCTTCTCTCA 416
Db      175 ACTTGCTGAGGACGATATCTGATGATGCTGCGGAGAGCGCCCATCTCATTTATTCATCA 234
Qy      417 ATTTGTAATCTTGTAGGGGCTCATCATGATGATGATGATGATGATGATGATGATGATG 476
Db      235 GTTATGATTTGCTGTGGGTTGATTAATGATGATGATGATGATGATGATGATGATGATG 294
Qy      477 TTGCAAGCTTGTGATGATGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 536
Db      295 TTGCTAAGCTTCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 354
Qy      537 ATTTCAAAATAGCCCTTGTGAGATTAAGATTTGCTGAATACATCAACCAATTCAGT 596
Db      355 ATTTCTGAACCGCTCTCCGAGATCAGAGGACAGTAAATATCTCTCCCTCAGTTTCTT 414

```


QY 597 GGATCAGAGGAGATGTTCTTGTCACTGATGATGTTGGATGTCCTGTGCGCATCA 656
 DB 415 GGCCTGGGAGATGTTTGTGATACATGATGTTTCACTATGTCCTGTGATGATCC 474
 QY 657 CCCGATTGAGAAATATGCTGTGTGTCGCGAATCACTTCAATGTTCTTCTTGGTTTG 716
 DB 475 CCTACTGAGAGACCATGCTGGTGTCTCTGATGATGCTTCTTCTTATTTGTTCTC 534
 QY 717 ACAATATTTTATCTTCTGATATCTCCAGATGCTCTGTTAGCAAGATGCGATGCGAG 776
 DB 535 AGGGTATTTATTTGCTCCAGATCTCTCGTTGGCTGTGATGATGAAAGAAATGACGAG 594
 QY 777 GCAAAAGAGTGTGCAAAAGTTAAGGGGAAAGAGATGCTCAAGTGAATTTGCCCT 836
 DB 595 GCTAAGCAGTCTTCAACAGTATGTCAGAGAAAGATGATCCATGATGATGCTTTA 654
 QY 837 CTCTCGAAGGTTGAGGTTGAGAGACACTTCCATTAAGAGATCAATGGAACCT 896
 DB 655 CTAGTTGAGAGACTGATATAGAGAGAGAAAGATGAGATCTTATGATCTTGG 714
 QY 897 GCCACGAGGACCGCATGAT--CTTGTACTGACGATGATAGAACAAATCACAATT 953
 DB 715 GAGGATTCATGAGATGATGATACACTTGAACCGTTGATGAGATGACAAATGCGGCTT 774
 QY 954 TATGGGCTGGAAGAGCCAGTCAATGATGCTGCACTTTAAGGACCCATCAATGCTT 1013
 DB 775 TATGGAACCCAGAGATCAATGATGATCTTGTAGCTGATCCAGAACAA-- 826
 QY 1014 GGAAGTGTCTTCTTCTGATCTGTCATGAGAGATGATGATGACAGATGATCCCTT 1073
 DB 827 -ATAGTCACTTGGGCTAGGCTGTGCGCAAGGAGTTAGCAAGCAAGATGATCTT 885
 QY 1074 ATGATTCGATGATGATCACTTTTGTGATGTCATGAGATATGCTCAAGCTGAGGA 1133
 DB 886 AAGATATCCGCTGTCAATCTTTTGTGAGCTTCCAGAGATGTCAGAAAGAGCGGCA 945
 QY 1134 AGTATGAGAGACATGTTTCCAACTTTGGAAGATGATGAG--TGTCACAGAT 1187
 DB 946 AACACTCGGAGTGGATTTTCCCTCATTTGGAAGATGTTCACTGATCTGCGATGCG 1005
 QY 1188 CAGCATGCAAAATATGAGATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGATGAC 1238
 DB 1006 CCTCAAGGTAACCGGCTCATTTGAGAGAGAGATGAGAGAGATGAGAGAGAGAT 1065
 QY 1239 GAGGATGATGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1289
 DB 1066 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1125
 QY 1290 AGCCATGCTGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGAT 1349
 DB 1126 AGCCCTTATGATGTCGCGCAAGACACAGAGATGAGAGAGAGATGATGATGATGAT 1182
 QY 1350 CACCGTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1409
 DB 1183 AAGATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1239
 QY 1410 GTGAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1469
 DB 1240 GAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1279
 QY 1470 GGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1529
 DB 1280 -----ATGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1529
 QY 1530 CCTGGCTCAAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1589
 DB 1330 GCTGATATCTGCGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1383
 QY 1590 AGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1649
 DB 1384 GCGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1443
 QY 1650 GGTGAGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1709

DB 1444 CATGGA-----TCGACATGATGTTCCCGGAGAGAAATGCTGCTCTGGA 1488
 QY 1710 TCACGTTGAGAAATTTGTTGAACCTGAGTAGAGCGTCCCTGTTAGTCGGTTGGA 1769
 DB 1489 CCACCTGCTGCTCTTCTTGAACCTGTTGAAGCGTCCCTGTTAGTGTGTTGCTGAGC 1548
 QY 1770 ATTGATATCTTCAACAGATTTGCTGATATTAAGGTTGTTGTTGTTGTTGTTGTTGTTGTT 1829
 DB 1549 ATTCAAAATACGACAGATTTTCAAGTATGATGATGATGATGATGATGATGATGATGAT 1608
 QY 1830 CTTGAGAGCTGTGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1889
 DB 1609 CTCGAGAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1668
 QY 1890 TCCATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1949
 DB 1669 TCATCTCTCATGAGGCTTAAACATTAATGATGATGATGATGATGATGATGATGATGATGAT 1728
 QY 1950 CTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2009
 DB 1729 AGACTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1788
 QY 2010 GCATCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2069
 DB 1789 GTCTACTGTTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1848
 QY 2070 TTGCTCTCACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2129
 DB 1849 GCATCTCTCACAGGTTGTTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1908
 QY 2130 ATCCCAACATTTATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2189
 DB 1909 ATTCAAAATCTGTTGTTGAAGTTTCCAAAGAGTCCGTTGCTGATGATGATGATGAT 1968
 QY 2190 ATTGTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2249
 DB 1969 ATATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2028
 QY 2250 CTGAATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2309
 DB 2029 CTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2088
 QY 2310 TTTGTTGCTGCTTCTTAAAGTCCCTGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2369
 DB 2089 TGGATCTTCTGTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2148
 QY 2370 GAATCTTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2389
 DB 2149 GACTACTTTCCTTTCGAGC 2168
 RESULT 7
 ABKS1968
 ID ABKS1968 standard; cDNA; 1487 BP.
 XX ABKS1968;
 XX 27-AUG-2002 (first entry)
 DE Wheat contig encoding Arabidopsis thaliana-like sugar transport protein.
 XX
 AC Wheat; Arabidopsis thaliana-like sugar transport protein;
 KW carbohydrate transport; grain filling; annual field crop;
 XX plant; gene; ss.
 OS Triticum aestivum.
 XX
 XX Key Location/Qualifiers
 FH 3..1040
 FT /*tag= a
 FT /partial
 FT /product= "Portion of a wheat Arabidopsis

XX O
 DR MPI: 2003-340957/32.
 DR P-PSDB; AB008332.
 XX
 XX Novel plant sugar transport proteins and nucleic acid encoding the
 PT protein useful for producing transgenic plants having altered levels of
 PT sugar transport protein -
 PS Claim 2; Page 25-26; 56pp; English.
 XX
 XX The present invention relates to the isolation of Arabidopsis
 CC thaliana-like or Beta vulgaris-like sugar transport proteins, and
 CC the polynucleotide sequences encoding them. The plant sugar transport
 CC proteins of the invention have been isolated from corn, rice, soybean,
 CC and wheat. The polypeptides of the invention may be used for altering
 CC the level of expression of a sugar transport protein in a host cell,
 CC by transforming a host cell with a chimeric construct encoding all,
 CC or a portion of the sugar transport protein, in sense or antisense
 CC orientation. Particularly, the polypeptides may provide a means to
 CC control carbohydrate transport and distribution in plants.
 CC ABX9198-ABX93205 represent cDNA sequences encoding Arabidopsis
 CC thaliana-like sugar transport proteins.
 SQ Sequence 1487 BP; 355 A; 320 C; 373 G; 439 T; 0 other;
 Query Match 24.9%; Score 691.2; DB 25; Length 1487;
 Best Local Similarity 80.1%; Pred. No. 1.3e-153;
 Matches 852; Conservative 0; Mismatches 203; Indels 9; Gaps 3;

QY 1395 GAGGGTGAGATGTGTGAGCAGCACTGATCGTGGGGGATGCGAGCTTGCGA 1454
 DB 12 GAGGGTGGGAGGAGCACTGAGCACTGATTTGGGGGGGAGCACTCGCATGAAA 71
 QY 1455 TGGTCAG 1514
 DB 72 TGGTCGAG 131
 QY 1515 CACCAAG 1574
 DB 132 CACCAAG 191
 QY 1575 GATG---TTTTGAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1631
 DB 192 GATGCCAG 251
 QY 1632 GTTTCTCAAG 1688
 DB 252 GTTTCTCAAG 311
 QY 1689 TCTGAG 1748
 DB 312 TTGGAG 371
 QY 1749 GCGCTGTGAG 1808
 DB 372 GCGCTGTGAG 431
 QY 1809 CTGTACTATATCCCAAAATTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1868
 DB 432 CTGTACTATATCTCTCAAAATTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 491
 QY 1869 GGTCTCAG 1928
 DB 492 GGTCTCAG 551
 QY 1929 CCTGAG 1988
 DB 552 CCTGAG 611
 QY 1989 GGCAG 2048
 DB 612 GGCAG 671

QY 2049 TTGGGTACAG 2108
 DB 672 TTGGGTACAG 731
 QY 2109 TTGGGTACAG 2168
 DB 732 TTGGGTACAG 791
 QY 2169 GTTGTGAG 2228
 DB 792 GTTGTGAG 851
 QY 2229 ACCTACAG 2288
 DB 852 ACCTACAG 911
 QY 2289 GCGAGTGTGAG 2348
 DB 912 GCGAGTGTGAG 971
 QY 2349 ATGCCCCCTGAG 2405
 DB 972 ATGCCCCCTGAG 1031
 QY 2406 AAGGCTAATTTCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2449
 DB 1032 AAGGCTAATTTCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1075

RESULT 9
 ABK51966
 ID ABK51966 standard; cDNA, 1692 BP.
 XX
 XX ABK51966;
 XX
 XX 27-AUG-2002 (first entry)
 XX
 DE Soybean cDNA clone ssl.pk0022.f1 encoding sugar transport protein.
 XX
 KW Soybean; Arabidopsis thaliana-like sugar transport protein;
 KW carbohydrate transport; grain filling; annual field crop;
 KW plant; clone ssl.pk0022.f1; gene; ss.
 XX
 OS Glycine max.
 XX
 XX Key Location/Qualifiers
 FH 9..1469
 FT /*tag= a
 FT /product= "Portion of a soybean Arabidopsis
 FT thaliana-like sugar transport protein"
 FT /note= "This sequence lacks a start codon"
 XX
 PN US6383776-B1.
 XX
 PD 07-MAY-2002.
 XX
 PF 14-APR-1999; 99US-0291922.
 XX
 PR 24-APR-1998; 98US-083044P.
 XX
 PA (DUPO) DU POINT DE NEMOURS & CO E I.
 PI Allen SM, Hiltz WD, Kinney AJ, Tingey SV;
 XX
 DR MPI: 2002-451386/48.
 DR P-PSDB; AAU97205.
 XX
 XX New nucleic acid encoding plant sugar-transport proteins, useful for
 PT preparing transgenic plants with altered carbohydrate distribution
 XX
 PS Claim 3; Column 41-44; 54pp; English.
 XX

CC The present invention relates to the isolation of plant polynucleotide
 CC sequences encoding an Arabidopsis thaliana-like sugar transport
 CC protein or Beta vulgaris-like sugar transport protein. The
 CC polynucleotide sequences are useful for altering the level of sugar
 CC transport proteins in plants, i.e. for control of carbohydrate transport
 CC and distribution in plant cells, e.g. during grain filling of annual
 CC field crops (e.g. corn, rice, soybeans, and wheat), and, for studying
 CC carbohydrate flows and sugar transport. The polynucleotide
 CC sequences can also be used to isolate cDNA sequences and genes that
 CC encode homologues of the new proteins. The present sequence encodes
 CC a portion of a soybean Arabidopsis thaliana-like sugar transport
 CC protein.

XX Sequence 1692 BP; 430 A; 339 C; 400 G; 523 T; 0 other;

Query Match 21.4%; Score 594.6; DB 24; Length 1692;

Best Local Similarity 66.1%; Pred. No. 1e-130; Mismatches 459; Indels 39; Gaps 6;

Matches 969; Conservative 0; Mismatches 459; Indels 39; Gaps 6;

QY 929 CGGTGTAAGAAACAATCACTTATGAGGCTGAGAGAGCCAGTCATGATGCTCG 988
 DB 17 CAGAGAAAAGATCAATTAAGTGTATGACCAAGAACAGGCACTCTGGTCTGAG 76
 QY 989 ACCCTTAAGGACCCATCATGCTTGAAGTGTCTTCTTGTGATCTGTCAGGAG 1048
 DB 77 ACCCTTGTCTGACCAATTCCTGTTGTC-----CTTGTATCTAGGAAAGAG 124
 QY 1049 CATTGGAACAGAGGTACCCCTTATGATCCGATTGAGACCTTTTGTATGTCGA 1108
 DB 125 CATTGGAACAGAGGTACCCCTTATGATCCGATTGAGACCTTTTGTATGTCGA 181
 QY 1109 TGAAGATATGCTCAAGCTGAGAGAGATGAGAGACATGTTTCCAACTTTGAG 1168
 DB 182 TGAAGATATGCTCAAGCTGAGAGAGATGAGAGACATGTTTCCAACTTTGAG 229
 QY 1169 TATGTCAGTGTCAAGATCAAGTCAAGTCCAAATGAGAGAGTGAAGAGATCTTA 1228
 DB 230 TATGTCAGTGTGAGAGATCAAGTCAAGTGAAGATGAGAGAGATGAGAGAG 289
 QY 1229 TGGAGTACAGAGATGAGATGATGATGAGAGAGTGAAGATGAGAGATGAG 1288
 DB 290 CAGAGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 343
 QY 1289 TAGCCATTGCTGTCCAGGAGAGCAAGGTGAGAGAGAGAGAGATGAGAGATG 1348
 DB 344 GAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 400
 QY 1349 TCACCGTGAAGTCTTTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1408
 DB 401 CCAATGATTAACCTTGCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 460
 QY 1409 TGTGACACACATGATGATGATGATGATGATGATGATGATGATGATGATG 1468
 DB 461 CACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 520
 QY 1469 AGGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1528
 DB 521 GGGCCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 580
 QY 1529 TCTGCTCAAG 1588
 DB 581 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 637
 QY 1589 TAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1648
 DB 638 CAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 697
 QY 1649 TGCTGAACACGATGATGATGATGATGATGATGATGATGATGATGATGATG 1708
 DB 698 TATGCTCAAG 757
 QY 1709 TTCAGTGAAG 1768

DB 758 GCCAAGTGAAGATCTTTTGAACCTGGGAGAGAGATGATGATGATGATGATG 817
 QY 1769 AATTGAGATCTTCAACAGTTTGGAGATGAATAAGAGGTGTTGATGATGATGAT 1828
 DB 818 AATTGAGATCTTCAACAGTTTGGAGATGAATAAGAGGTGTTGATGATGATGAT 877
 QY 1829 TCTTGAAG 1888
 DB 878 TCTTGAAG 937
 QY 1889 ATTCATCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 1948
 DB 938 ATTCATCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 997
 QY 1949 GCTGCTTATGATCTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2008
 DB 998 GAGGCTCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1057
 QY 2009 AGCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2068
 DB 1058 AGCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1117
 QY 2069 TTTGCTTCAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2128
 DB 1118 ATCAATCTCAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1177
 QY 2129 CATCCCAACATTTATGAGAGAGATCTTCCACAGAGAGAGAGAGAGAGAG 2188
 DB 1178 ATTCATCTCAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1237
 QY 2189 CATTGTCCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2248
 DB 1238 TATTTGTCCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1297
 QY 2249 GCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2308
 DB 1298 GCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1357
 QY 2309 CTTTGTGTCCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 2368
 DB 1358 ATGGAGTGTGCTTTTGAAGAGTTCAGAAACAGAGAGAGAGAGAGAGATG 1417
 QY 2369 CGAATCTTTCAGTGTGTCAGAGCA 2395
 DB 1418 TGAATCTTCTGTCGAGAGCAAAACA 1444

RESULT 10
 ABX93202
 ID ABX93202 standard; cDNA; 1692 BP.
 AC ABX93202;
 DT 29-MAY-2003 (first entry)
 XX
 DE cDNA encoding soybean sugar transport protein #2.
 KW Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
 KW plant sugar transport protein; carbohydrate transport; soybean;
 KW carbohydrate distribution; plant; gene; ss.
 OS Glycine max.
 XX
 PN US2002178468-A1.
 PD 28-NOV-2002.
 XX
 PF 17-JAN-2002; 2002US-0051902.
 XX
 PR 24-APR-1998; 98US-083044P.
 PR 14-APR-1999; 99US-0291922.
 XX
 PA (ALLEN/) ALLEN S M.

KW		carbohydrate transport; grain filling; annual field crop;
KW		plant; clone wreln.pk0006.b4; gene; ss.
OS	Triticum aestivum.	
XX		
XX	Key	Location/Qualifiers
FT	CDS	2..688
FT		/+tag= a
FT		/partial
FT		/product= "Portion of a wheat Arabidopsis thaliana-like sugar transport protein"
FT		/note= "This sequence lacks a start codon"
XX		
PN	US6383776-B1.	
PD	07-MAY-2002.	
XX		
PF	14-APR-1999;	99US-0291922.
PR	24-APR-1998;	98US-083044P.
PA	(DUPO) DU PONT DE NEMOURS & CO E I.	
PI	Allen SM, Hltz WD, Kinney AJ, Tingey SV;	
DR	WPI; 2002-451386/48.	
DR	P-PSDB; AAU97208.	
XX		
PT	New nucleic acid encoding plant sugar-transport proteins, useful for preparing transgenic plants with altered carbohydrate distribution -	
XX	Claim 3; Column 51-54; 54pp; English.	
XX		
CC	The present invention relates to the isolation of plant polynucleotide sequences encoding an Arabidopsis thaliana-like sugar transport protein or Beta vulgaris-like sugar transport protein. The polynucleotide sequences are useful for altering the level of sugar transport proteins in plants, i.e. for control of carbohydrate transport and distribution in plant cells, e.g. during grain filling of annual field crops (e.g. corn, rice, soybeans, and wheat), and, for studying carbohydrate flows and sugar transport. The polynucleotide sequences can also be used to isolate cDNA sequences and genes that encode homologues of the new proteins. The present sequence encodes a portion of a wheat Arabidopsis thaliana-like sugar transport protein.	
SQ	Sequence 1009 BP; 237 A; 233 C; 234 G; 305 T; 0 other;	
	Query Match	14.2%; Score 394.4; DB 24; Length 1009;
	Beech Local Similarity	74.4%; Fred. No. 2.3e-83;
	Matches 497; Conservative	0; Mismatches 171; Indels 0; Gaps 0;
OY	1730	TGAACCTGAGTAGAGCGTCGCCGTGGTATGTCGGTGTTGGAAATGCAGATCCTTCAACAGTT 1789
Db	1	TGAACCTGAGTAGAGCATGACATGCTTGTGGCATAGAGATTACAGATCTTCGACAGATT 60
OY	1790	TGCTGGAATAAACGGTGTTCTGTACTATATACCACAATAATTTGAGCAAGCTGTGTGC 1849
Db	61	TGCGGGATCAATGAGATGATGCTCTACTACACACTCAATTAATTTAGGAGAAGGATGTCGG 120
OY	1850	AGTTATTTTTCCAAATTTGCTTCAGCTGGGATCAGATTCATCTTGATCAAGTTCTCT 1909
Db	121	GGTCTTCTTATCAAAACATTTGACATGACTCTTCCGACATCTAATTTTATTAGGCTT 180
OY	1910	CACATACCTTAAATAGCTTCCTGATTTGGCTTTGGATTCAGTATAGATCTTTCCG 1969
Db	181	GACACCTTCTGATGATGCTTCCAGCATTTGGATGCGCATGATTCATGAGATATATGACG 240
OY	1970	AAGAAGTTTTTGTCTGCTAGGACAAATTCATCTTGATAGCATCTAGTTATCTGGT 2029
Db	241	AAGAAGTTTTCTTCTCTTTCAACATCCCTGTCTTGATAGAGGCTAGCTGTCTGGT 300
OY	2030	TGTGTCAATCTAATGATTTGGGTGACCTAGCCATGCTTGTCTCTCACCGTCAGGT 2089

Db	301	TTTAGTGAAGTCTGGATGTCGGAAACATGATGACGCTGCGCTCTCAACGATCAGCGT	360
Oy	2090	TATCGTCTACTTCTGCTGCTCTGTTATGAGATTGGTCCCATCCCAACATTTATGTGC	2149
Db	361	CATGCTCATTTCTGCTTCTTCGTCATGAGGGTTTGAGGCTATCCCAATATATTCCTGCGC	420
Oy	2150	AGACATCTTTCACCAACGAGGTTGATGAGCTGATGATGATGCAATTTGACCTTACATCTG	2209
Db	421	GGAAATTTTTCCTCCCACTCTGTCGTCATTCGATAGCCATCTGCGGCTAACCTTCTG	480
Oy	2210	GATCGGAGATATCATCTGTCACCTTCACACCTTCCGTATGCTGAATCTTATGACTGGC	2269
Db	481	GATGGGGACATCATCTGTCATATCACTCTCCCGTATGTCATATGCCATTTGGTCTCGC	540
Oy	2270	GGGTGTTTTCAGCATATATGACATGCTGATGCTGATTTCTTTGTTGTTGCTCTTCCTTAA	2329
Db	541	TGGAGTCTTTCGGCATATATGSCATCTGTTGTGATGACTGACCTTTGTATTCGTCATGAA	600
Oy	2330	GGTCCCTGAGACAAAGGGATGCCCCCTTGAGGTTATATACGAATTCCTTTCAGTTGGTGC	2389
Db	601	GGTCCCTGAGACAAAGGGATGCCCCCTGAGGTCATCACGAGTTCTTCTCTGCGGAGC	660
Oy	2390	GAAGCAAG 2397	
Db	661	AAAGCAGG 668	
RESULT 12			
ID	ABX93205		
AC	ABX93205 standard; cDNA; 1009 BP.		
XX	ABX93205;		
XX			
XX	29-MAY-2003 (first entry)		
DE	cDNA encoding wheat sugar transport protein #3.		
XX			
KW	Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;		
KM	plant sugar transport protein; carbohydrate transport; soybean;		
KW	carbohydrate distribution; plant; gene; ss.		
XX			
OS	Triticum aestivum.		
XX			
PN	US2002178468-A1.		
PD			
XX	28-NOV-2002.		
PF	17-JAN-2002; 2002US-0051902.		
XX			
XX	14-APR-1998; 98US-083044P.		
PR	14-APR-1999; 99US-0291922.		
XX			
PA	(ALLEN/) ALLEN S. M.		
PA	(HITZ/) HITZ W. D.		
PA	(KINNEY/) KINNEY A. J.		
PA	(TINGE/) TINGEY S. V.		
XX			
PI	Allen SM, Hiltz WD, Kinney AJ, Tingey SV;		
XX			
DR	WPI; 2003-340957/32.		
DR	P-PSDB; ABU08333.		
XX			
PT	Novel plant sugar transport proteins and nucleic acid encoding the		
PT	protein useful for producing transgenic plants having altered levels of		
PT	sugar transport protein -		
XX			
PS	Claim 2; Page 27; 56pp; English.		
XX			
CC	The present invention relates to the isolation of Arabidopsis		
CC	thaliana-like or Beta vulgaris-like sugar transport proteins, and		
CC	the polynucleotide sequences encoding them. The plant sugar transport		
CC	proteins of the invention have been isolated from corn, rice, soybean,		
CC	and wheat. The polypeptides of the invention may be used for altering		

CC the level of expression of a sugar transport protein in a host cell,
 CC by transforming a host cell with a chimeric construct encoding all,
 CC or a portion of the sugar transport protein, in sense or antisense
 CC orientation. Particularly, the polypeptides may provide a means to
 CC control carbohydrate transport and distribution in plants.
 CC ABX3198-ABX3205 represent cDNA sequences encoding Arabidopsis
 CC thaliana-like sugar transport proteins.

XX Sequence 1009 BP, 237 A, 233 C, 234 G, 305 T, 0 other;

Query Match 14.2%; Score 394.4; DB 25; Length 1009;
 Best Local Similarity 74.4%; Pred. No. 2.3e-63;
 Matches 497; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY 1730 TGAACCTGAGTGGAGGCGCTGCTTAACTCGTGTGGAATTCATCTTCAACAGTT 1789
 DB 1 TGAACCTGAGTGGAGGCGCTGCTTAACTCGTGTGGAATTCATCTTCAACAGTT 60
 QY 1790 TGGTGAATAAACGGTCTCTGACTATACCCAGAAATTTGAGCAAGCTGGTGGC 1849
 DB 61 TGGGATATCAATGAGTCTCTACTACACACTCGATCTTGAGCAAGAGGTGTGG 120
 QY 1850 AGTATATCTTCCAAATTTGATCTCAGCTCGGATCAGCATCCTGATCTGATCTCT 1909
 DB 121 GGTCTTCTATCAACATTTGAGTGAAGCTCTTCTCAGCATCTATCTTATAGTCTT 180
 QY 1910 CACTACCTTACTATGCTTCTTGGCATTTGGCTTTGCGATGCTCTTATGAGATCTTCCG 1969
 DB 181 GAGAACCTTGTGATGCTTCCGACATTTGCGATGCGCATGAGATCATGATATGTCAG 240
 QY 1970 AAGAAAGTTTGTGCTGCTAGGACATTTGCAATCTTATGAGATCTCTGATATCTGCT 2029
 DB 241 AAGAAAGTTTGTGCTTCTTCAACATCTTCTGATATGAGTGAAGCTTGTGCTTGT 300
 QY 2030 TGTGCTCAATCTTATGATTTGGGTACACTAGCCCATGCTTGTCTCCACGCTCAGT 2089
 DB 301 TTTAGTGAATGTTTGGATGTTGGAGCATGAGTGCAGCGTGGCTTCAACGATAGCGT 360
 QY 2090 TATGCTTACTTCTGCTGCTTGTGTTATGAGATTTGCTCCATCCCAATTTATGCT 2149
 DB 361 CATGCTTATTTCTGCTTCTGCTCATGAGGAGTTGGGCTTATCCCAATTTCTGCGC 420
 QY 2150 AAGAAATCTTCAACAGGAGTGTGAGCTCTGATTTGCAATTTGCTTATCAATCTG 2209
 DB 421 GGAATATTTTCCCACTCTGCTGCTGAGTGCATGAGCATCTGCGCTTAACTTCTG 480
 QY 2210 GATCGAGATATCATGCTCAGCTACAGCTTCTCTGATGCTGAATGATTTGAGTGGC 2269
 DB 481 GATCGGCGATCATGCTGATCATACCTCTCCCGATGATGCTCAATGATGCTGCTGC 540
 QY 2270 GGGTGTTCAGCATATATGACAGTGTGATGCTTGAATTTCTTGTGCTTCTTAA 2329
 DB 541 TCGAGTCTTCCGATATATGATCTTGTGATCTGATGCTTGTATGCTTACATGAA 600
 QY 2330 GGTCCCTGAGCAAGAGGAGTGGCTTGAAGTATATACGATTTTGGAGTGGTGC 2389
 DB 601 GGTCCCTGAGCAAGAGGAGTGGCTTGAAGTATATACGATTTTGGAGTGGTGC 660
 QY 2390 GAAGCAAG 2397
 DB 661 AAGGACAG 668
 RESULT 13
 ABL71325
 ID ABL71325 standard; cDNA; 282 BP.
 AC ABL71325;
 XX 14-MAY-2002 (first entry)
 DT
 XX
 DE Corn tassels-derived polynucleotide (cdps) SEQ ID NO:699.
 XX

KW Corn: corn tassels-derived polynucleotide; cdps; hybrid breeding; CDPS;
 KW inheritance; characteristic; growth; development; disease resistance;
 KW environmental adaptability; quality; yield; molecular marker;
 KW multigene trait; plant breeding; corn tassels; gene; ss.
 OS Zea mays.

XX US2001051335-A1.

XX 13-DEC-2001.

XX 16-APR-1999; 99US-0294093.

XX 21-APR-1998; 98US-082567P.

XX (LALG/) LALGUDI R V.

XX (ITOL/) ITO L Y.

XX (SHER/) SHERMAN B K.

XX Lalgudi RV, Ito LY, Sherman BK;

XX WPI; 2002-163647/21.

PT Novel purified corn tassels-derived polynucleotide useful for
 PT determining altered gene expression, to recover regulatory elements and
 PT to follow inheritance of desirable characteristics through hybrid
 PT breeding programs -

PS Claim 1; SEQ ID 699; 201pp; English.

CC The present sequence describes a purified corn tassels-derived
 CC polynucleotide sequence (cdps) comprising a nucleic acid sequence
 CC selected from those given in ABL70627 to ABL76633. The cdps sequences
 CC encode corn tassels-derived polypeptides (CDPS). The cdps sequences (I)
 CC can be used for determining altered gene expression, to recover
 CC regulatory elements and to follow inheritance of desirable
 CC characteristics through hybrid breeding programs. (I) are also useful
 CC in the evaluation, and alteration of desired characteristics associated
 CC with growth and development, disease resistance, environmental
 CC adaptability, quality and yield, and as molecular markers for studying
 CC inheritance of multigene traits in a plant breeding program. (I) can be
 CC used to produce a tassels-specific profile of gene transcription, a
 CC transcript image, to clone regulatory elements for use in transformation
 CC vectors, to express a polypeptide, to identify, isolate or extend
 CC identical or related corn tassels nucleic acid sequences from DNA
 CC libraries, in nucleic acid hybridisation or amplification technologies,
 CC as query sequences to determine homology of known sequences, as probe
 CC for use in Southern or Northern hybridisation, and to identify the
 CC presence of and/or to determine the degree of similarity between two
 CC (or more) nucleic acid sequences.

XX Sequence 282 BP, 54 A, 63 C, 61 G, 95 T, 9 other;

Query Match 9.3%; Score 257.2; DB 24; Length 282;
 Best Local Similarity 95.4%; Pred. No. 4.7e-51;
 Matches 270; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 1938 GGGTTGCGATGCTGCTTATGATCTTTCCGAGAGAGCTTTTGTCTGAGCAAT 1997
 DB 1 GGGTTGCGATGCTGCTTATGATCTTTCCGAGAGAGCTTTTGTCTGAGCAAT 60
 QY 1998 CCAATCTGATAGCATCTCTAGTATCTGCTGTTGTCATCAATGATTTGGGTACA 2057
 DB 61 CCAATCTGATAGCATCTCTAGTATCTGCTGTTGTCATCAATGATTTGGGTACA 120
 QY 2058 CTAGCCCATGCTTGTCTCTCAGCGTGAATCTCTACTTCTGCTGCTTCTATG 2117
 DB 121 CTAGCCCATGCTTGTCTCTCAGCGTGAATCTCTACTTCTGCTGCTTCTATG 180
 QY 2118 GGATTTGCTCCATCCCAACATTTTATGTCAGAGATCTTTCAACAGGAGTGTG 2177
 DB 181 GGATTTGCTCCATCCCAACATTTTATGTCAGAGATCTTTCAACAGGAGTGTG 240

QY 2178 CTCGTGATTCGATTTGGCCCTTACATTCGTGATCGAGATA 2220
 |||||
 DB 241 CTCGTGATTCGATTTGGCCCTTACATTCGTGATCGAGATA 282
 |||||
 RESULT 14
 ID ABR51964
 XX ABR51964 standard; cDNA; 870 BP.
 AC ABR51964;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE Rice cDNA clone rl56.pk0003.d5 encoding sugar transport protein.
 XX
 KW Rice; Arabidopsis thaliana-like sugar transport protein;
 KM carbohydrate transport; grain filling; annual field crop;
 XX plant; clone rl56.pk0003.d5; gene; ss.
 OS Oryza sativa.
 XX
 PH Key Location/Qualifiers
 FT CDS 8..403
 FT /+tag= a
 FT /partial
 FT /product= "Portion of a rice Arabidopsis
 thaliana-like sugar transport protein"
 FT /note= "This sequence lacks a start codon"
 XX
 PN US6383776-B1.
 XX
 PD 07-MAY-2002.
 XX
 PF 14-APR-1999; 99US-0291922.
 XX
 PR 24-APR-1999; 98US-083044P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E. I.
 XX
 PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
 XX
 DR WPI: 2002-451386/48.
 XX
 DR P-PSDB; AAB97203.
 XX
 PT New nucleic acid encoding plant sugar-transport proteins, useful for
 preparing transgenic plants with altered carbohydrate distribution
 XX
 PS Claim 3; Column 33-36; 54pp; English.
 XX
 CC The present invention relates to the isolation of plant polynucleotide
 sequences encoding an Arabidopsis thaliana-like sugar transport
 protein or Beta vulgaris-like sugar transport protein. The
 CC polynucleotide sequences are useful for altering the level of sugar
 transport proteins in plants, i.e. for control of carbohydrate transport
 CC and distribution in plant cells, e.g. during grain filling of annual
 CC field crops (e.g. corn, rice, soybeans, and wheat), and for studying
 CC carbohydrate flows and sugar transport. The polynucleotide
 CC sequences can also be used to isolate cDNA sequences and genes that
 CC encode homologues of the new proteins. The present sequence encodes
 CC a portion of a rice Arabidopsis thaliana-like sugar transport protein.
 XX
 SQ Sequence 870 BP; 191 A; 186 C; 206 G; 287 T; 0 other;
 Query Match 8.2%; Score 227; DB 24; Length 870;
 Best Local Similarity 74.9%; Pred. No. 1e-43;
 Matches 284; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
 QY 2020 TTATCTGATTCGATTCATTAATGATTTGGGTACACGCGCATGCTTGTCTTCA 2079
 |||||
 DB 12 TAACTTGTATTCGTCATATTCATTCGATGCGGACCATGCTTCACTACCTTCA 71
 |||||
 QY 2080 CCGTCAGTGTATTCGTCATTCCTTGTGCTTCGTTAATGGGATTTGGTCCCATCCCA 2139
 |||||

DB 72 CAGTCAGTCATACCTACTTCCTCTTGTGATGAGGAGTTCGAGCTATTCACAA 131
 QY 2140 TTTATATGCGAGATCTTTCCACAGAGTTGCGCTGTATTCGATTTGCT 2199
 |||||
 DB 132 TTCTCTGCGAGAGATTTTCCGACACCGTTGCGATCTGCATACCATCTGTGCCC 191
 |||||
 QY 2200 TTACATTCGATTCGAGATATCATCTGCACCTACAGCTTCTGTGATGTAATGCTA 2259
 |||||
 DB 192 TAACTTGTGATTCGTCATATTCATCTGACATACACCTCCCGGTGATGCTCAAGCCA 251
 |||||
 QY 2260 TTGACCTGCGCGGTGTTTCAGCATATATGACATGCTGATTTCTTGTGCTG 2319
 |||||
 DB 252 TTGACCTGCGCTGAGATGTTTGAATCTACGACGTGCTGCATACCTGCTTCTGTTG 311
 |||||
 QY 2320 TCTTCTTAAGTCCCTGAGCAAGGAGGAGCCCTGAGTTATTAACCAATCTTGG 2379
 |||||
 DB 312 TCTTATGATGAGTTCGCGAGAACGAGGATGCTTGAATCATCAACGAGTCTTCT 371
 |||||
 QY 2380 CAGTTGTCGCAAGCAAGC 2398
 |||||
 DB 372 CTGTGAGCAAGCAAGCAGC 390
 |||||
 RESULT 15
 ID ABR93200
 XX ABR93200 standard; cDNA; 870 BP.
 AC ABR93200;
 XX
 DT 29-MAY-2003 (first entry)
 XX
 DE cDNA encoding rice sugar transport protein #2.
 XX
 KW Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
 KM plant sugar transport protein; carbohydrate transport; soybean;
 XX carbohydrate distribution; plant; gene; ss.
 OS Oryza sativa.
 XX
 PN US2002178468-A1.
 XX
 PD 28-NOV-2002.
 XX
 PF 17-JAN-2002; 2002US-0051902.
 XX
 PR 24-APR-1998; 98US-083044P.
 XX
 PR 14-APR-1999; 99US-0291922.
 XX
 PA (ALLEN) ALLEN S M.
 PA (HITZ) HITZ W D.
 PA (KINNEY) KINNEY A J.
 PA (TINGEY) TINGEY S V.
 XX
 PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
 XX
 DR WPI: 2003-140957/32.
 XX
 DR P-PSDB; AB008328.
 XX
 PT Novel plant sugar transport proteins and nucleic acid encoding the
 PT protein useful for producing transgenic plants having altered levels of
 PT sugar transport protein
 XX
 PS Claim 2; Page 18; 56pp; English.
 XX
 CC The present invention relates to the isolation of Arabidopsis
 CC thaliana-like or Beta vulgaris-like sugar transport proteins, and
 CC the polynucleotide sequences encoding them. The plant sugar transport
 CC proteins of the invention have been isolated from corn, rice, soybean,
 CC and wheat. The polypeptides of the invention may be used for altering
 CC the level of expression of a sugar transport protein in a host cell,
 CC by transforming a host cell with a chimeric construct encoding all,
 CC or a portion of the sugar transport protein, in sense or antisense
 CC orientation. Particularly, the polypeptides may provide a means to

CC control carbohydrate transport and distribution in plants.
CC ABX93126-ABX93205 represent cDNA sequences encoding Arabidopsis
CC thaliana-like sugar transport proteins.
XX
SQ Sequence 870 BP; 191 A; 186 C; 206 G; 287 T; 0 other;

Query Match 8.2%; Score 227; DB 25; Length 870;
Best Local Similarity 74.9%; Pred. No. 1e-43;
Matches 284; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY	2020	TTATCTGCTGTTGTCCTCAATCTAATGATTTGGGTACACTAGCCCATGCTTGTCTCCA	2079
DB	12	TAACTTGAATCTGTCATATTTCTGATGTGGGACATGTTCACTGCTCACTGTCCA	71
QY	2080	CCGTCACTGTTATGCTGTAATCTGCTGCTTGTATGGGATTTGGTCCCATGCCACA	2139
DB	72	CAGTCAGTGTCACTACTCTACTCTGCTTCTTGTGATGGGGTTGCGGGCTATTCANACA	131
QY	2140	TTTTATGTCAGAGATCTTCCAAACGAGGTTGTGGCCCTCTGTATTTGCCATTTGGCT	2199
DB	132	TTCTCTGTGTCAGAGATTTTCCGACACCGTTGTGTGTCATGTCATAGCCATCTGTGCC	191
QY	2200	TTACATTTCTGATCGGAGATATCATGTCACCTACAGCCTTCTGTGATGCTGAATGCTA	2259
DB	192	TAAATTTCTGATCGGTGATATCATGTCATACACCTCCCGATGCTCAACGCCA	251
QY	2260	TTGACTGCGCGGTGTTTTCAGCATATATGACGTCTGATTTCTTGTGTGG	2319
DB	252	TTGACTGCTGTGAGATTTTGGAACTAGCAGTGTGTCATGCTTGTCTGTGG	311
QY	2320	TCTTCTTAAGTCCCTGAGACAAAGGGATGCCCTTGAGTTATTACGAATTTCTTG	2379
DB	312	TCTTCATGAGGTGCGGAGACAAAGGGATGCTCTTGAAGTCATCACCAGTTCTCT	371
QY	2380	CAGTTGTGCGAAGCAAGC	2398
DB	372	CTGTGCGAGCAAGCAGGC	390

Search completed: January 5, 2004, 14:13:20
Job time : 746 secs

LOCATION: (896)
FEATURE:
NAME/KEY: unsure
LOCATION: (944)
US-09-291-922-1

Query Match 79.2%; Score 2200.2; DB 4; Length 2824;
Best Local Similarity 90.2%; Pred. No. 0;
Matches 2372; Conservative 0; Mismatches 253; Indels 6; Gaps 2;

133 CTTCACAGAGCGCGTACCTCGAGCATATCTTGGAGGACAGATGTCGGGGCTGTTTC 192
194 CTTGGCGCGCGCGCGAGATGGGGGCGTAGATTTCCGGGCGCATGGGGGGGCGCGTA 253
193 TTGTGCCCATAGTCGCTTCATGCGCAATCTATTGCGAGGGGGGGAACAATGCCACATG 252
254 TGGTCCCATCGCGGCGCTCATGCGCAACTTGTGCGAGGGCTGGGCAATGCGCAATATG 313
253 CAGCTGCTGTCTGTATATTAAGAGGAATTTCAATTTGCAAAATGAGCCCATGTCGAGAG 312
314 CTGAGACCGCTCGTACATTAAGAGGAATTTCAACTGCGAGCGAGCTTGATGAGAG 373
313 GACTAATTTGTCAATGTCATTTATCGCGCCACATCTGTTACTAATTTCCGGGCGCAT 372
374 GCGTCAATCGTCGCGCATGTTCTCATTTGGGGCAACATCATCAACAATCTCGGGGCGAA 433
373 TATCAACTGATGTCGCGCGACGCTATGCTTATTTCTCTTCAATTTGTAATTTCTTCA 432
434 GGGGCTGACTGCGTTGGTAGAGAGCCCATGCTGCTGCGCTGCTCTTCACTTCTGTA 493
433 GGGGCTCATCATGCTATGCTGCTATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 492
494 GTGGGCTGTGATGCTTTGGGCGCAATGTGTACATTTGCTCTCGCAAGGCTCATTTG 553
493 ATGATATTTGTAATTTGGCTTGTGCTGTCACGCTTGTGCTTGTATGATTTGAGAAATGACC 552
554 ATGGGTTGGTAATCGGTTTGGCGGTCAACTTGTCTCTCATCACTCGGAAATGCGAC 613
553 C--TTGGAATTAAGGTTTGTCTGAATACATCTACCAATTCAGTG--ATCAGAG 606
614 CCGACAGAAATCTTGGGCGCTGNTNGAACGTTGCGCGCATTCATTTGGGGGTCAAGAGAG 673
607 GAATGTTCTTGTATCTGATGCTGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 666
674 GGAATGTTCTCTCTACTGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 733
667 GAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 726
734 GGGTCAATGCTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 793
727 ATCTTCTGTAATCTCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 786
794 ACTTGTCTGAATCAACAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 853
787 TGTGCAAAAGTACGGGGGAAAGAGATGCTCAAGTGAATTTGCTCTTCTGAGAG 846
854 TGTGCAAAAGGCTGCGGGGAAAGAGATGCTCAAGGAGAGGCTCTTCTGAGAGAG 913
847 GGTGAGAGTTGAGAGAGACATTTCAATTTGAAGATGATCATTTGACCTGCCACGAGAG 906
914 GTTTGGGGGTGGTAAAGATACAGTATTTAGAGTACATTTGAGCTGCCACGAGAG 973
907 CAGCGCATGATCTTGTATCTGACGCTGATAGAGAAACAATCAGCTTTATGGGCTGAGAG 966
974 CAGCCCATATCTTGTATCTGACGCTGATAGAGAAACAATCAGCTTTATGGGCTGAGAG 1033
967 AAGGCGAGTATGATGCTGACGCTTTAAAGGAGCCCATGCTGAGAGAGGCTTT 1026
1034 AAGGCGAGTATGATGCTGACGCTTTAAAGGAGCCCATGCTGAGAGAGGCTTT 1093
1027 CTCTTGCATCTGCTCATGAGAGATGTAACAGAGTGAACCCCTTATGATCCGATTTG 1086
1094 CTCTTGCATCTGCTCATGAGAGATGTAACAGAGTGAACCCCTTATGATCCGATTTG 1153

1087 TGACACTTTTGTGATGCTCCATGAGATATGCTTCAAGCTTGAGAGAGATATGAGAGCA 1146
1154 TGACACTTTTGTGATGCTCCATGAGATATGCTTCAAGCTTGAGAGAGATATGAGAGCA 1213
1147 CATTTGTTCCAAACTTTGGAAGTATGTTCAAGTGTACAGATCAGATGCGCAAAAATGAGC 1206
1214 CATTTGTTCCAAACTTTGGAAGTATGTTCAAGTGTACAGATCAGATGCGCAAAAATGAGC 1273
1207 AGTGGAGTGAAGAGATCTTCAATGAGATGACGAGAGTACGATCTGATGTCAGAGAG 1266
1274 AGTGGAGTGAAGAGATCTTCAATGAGATGACGAGAGTACGATCTGATGTCAGAGAG 1333
1267 GTGACTTATGAGACATCTCCATTAACCATTTGCTCTCCAGGAGGCAACAAGTGGGAGAG 1326
1334 GTGACTTATGAGACATCTCCATTAACCATTTGCTCTCCAGGAGGCAACAAGTGGGAGAG 1393
1327 GGAAGGACATTTGACCATGCTGACCGTGAAGGCTTTGAGCATGAGAAAGCAAAACC 1386
1394 GGAAGGACATTTGACCATGCTGACCGTGAAGGCTTTGAGCATGAGAAAGCAAAACC 1453
1387 TCTTAGGGAGGGTGAAGATGCTGACGACGATATGCTGAGGGAGTGGACGCTTTG 1446
1454 TCTTAGGGAGGGTGAAGATGCTGACGACGATATGCTGAGGGAGTGGACGCTTTG 1513
1447 CTGGAATATGCTCAGAGAAAGGATGAGAAATGTAAGAAAGAAAGTGGTTTCAAAAAG 1506
1514 CTGGAATATGCTCAGAGAAAGGATGAGAAATGTAAGAAAGAAAGTGGTTTCAAAAAG 1573
1507 TCTACTGCAACCAAGAGGATTTCTGCTCAAGAGGGGCTCAATTTGTTCACTTCCCG 1566
1574 TCTACTGCAACCAAGAGGATTTCTGCTCAAGAGGGGCTCAATTTGTTCACTTCCCG 1633
1567 GTGAGGCGAGTGTGTAAGGGTATGAGATTTGTAATGCTGCTGCTGCTTATGATCACT 1626
1634 GTGAGGCGAGTGTGTAAGGGTATGAGATTTGTAATGCTGCTGCTGCTTATGATCACT 1693
1627 CAGCACTTTTCTCAAAAGGCTCTTGTGAACCGATGTCAGATGTCAGATGTCAGATGTC 1686
1694 CAGCACTTTTCTCAAAAGGCTCTTGTGAACCGATGTCAGATGTCAGATGTCAGATGTC 1753
1687 CATCTGAGGTAGCTGCCAAAGGATTCAGCTGGAAGATTTGTTGTAACCTGAGTGAAGC 1746
1754 CATCTGAGGTAGCTGCCAAAGGATTCAGCTGGAAGATTTGTTGTAACCTGAGTGAAGC 1813
1747 GTGCTTGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1806
1814 GTGCTTGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1873
1807 TTTCTGATCTAATCCCAAAATTTCTTGAAGCAAGCTGCTGCTGCTGCTGCTGCTGCTG 1866
1874 TTTCTGATCTAATCCCAAAATTTCTTGAAGCAAGCTGCTGCTGCTGCTGCTGCTGCTG 1933
1867 TTGCTCTGAGCTCGGATCAGCATCATCTTATGATGATCTCTCACTACTACTAATGTC 1926
1934 TTGCTCTGAGCTCGGATCAGCATCATCTTATGATGATCTCTCACTACTACTAATGTC 1993
1927 TTCTTGAATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1986
1994 TTCTTGAATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2053
1987 TTAGGCAAAATTCATCTTGAAGATCTTATGATCTTATGATCTTATGATCTTATGATCT 2046
2054 TTAGGCAAAATTCATCTTGAAGATCTTATGATCTTATGATCTTATGATCTTATGATCT 2113
2047 ATTTGGGTACATGACCATGCTTGTCTCTCAACGCTGATGTTATGCTACTTCTGCT 2106
2114 ATTTGGGTACATGACCATGCTTGTCTCTCAACGCTGATGTTATGCTACTTCTGCT 2173
2107 GCTTGTATATGAGATTTGTGCTCATCCCAACATTTATATGTCAGAGATCTTTCAACCA 2166
2174 GCTTGTATATGAGATTTGTGCTCATCCCAACATTTATATGTCAGAGATCTTTCAACCA 2233

QY 2167 GGGTTCGAGCCTCTGATATGCAATTTGACCTTACATTCGTGATCGAGATATCATCG 2226
DB 2234 GGGTTCGAGCCTCTGATATGCAATTTGACCTTACATTCGTGATCGAGATATCATCG 2293
QY 2227 TCACCTACAGCCTTCTGATATGCAATTTGACCTTACATTCGTGATCGAGATATCATCG 2286
DB 2294 TCACCTACAGCCTTCTGATATGCAATTTGACCTTACATTCGTGATCGAGATATCATCG 2353
QY 2287 ATGACGCTGATATGCAATTTGACCTTACATTCGTGATCGAGATATCATCG 2346
DB 2354 ATGACGCTGATATGCAATTTGACCTTACATTCGTGATCGAGATATCATCG 2413
QY 2347 GATGCTGATATGCAATTTGACCTTACATTCGTGATCGAGATATCATCG 2406
DB 2414 GATGCTGATATGCAATTTGACCTTACATTCGTGATCGAGATATCATCG 2473
QY 2407 ATGACGCTGATATGCAATTTGACCTTACATTCGTGATCGAGATATCATCG 2466
DB 2474 ATGACGCTGATATGCAATTTGACCTTACATTCGTGATCGAGATATCATCG 2533
QY 2467 GGGTTCGAGCCTCTGATATGCAATTTGACCTTACATTCGTGATCGAGATATCATCG 2526
DB 2534 GGGTTCGAGCCTCTGATATGCAATTTGACCTTACATTCGTGATCGAGATATCATCG 2593
QY 2527 ACATCTGATATGCAATTTGACCTTACATTCGTGATCGAGATATCATCG 2586
DB 2594 ACATCTGATATGCAATTTGACCTTACATTCGTGATCGAGATATCATCG 2653
QY 2587 ATGACGCTGATATGCAATTTGACCTTACATTCGTGATCGAGATATCATCG 2646
DB 2654 ATGACGCTGATATGCAATTTGACCTTACATTCGTGATCGAGATATCATCG 2713
QY 2647 ATGACGCTGATATGCAATTTGACCTTACATTCGTGATCGAGATATCATCG 2706
DB 2714 ATGACGCTGATATGCAATTTGACCTTACATTCGTGATCGAGATATCATCG 2773
QY 2707 TGTATATATATCAATCTCAATTAAGAAATATGTTTCTCAAAAAA 2757
DB 2774 TGTATATATATCAATCTCAATTAAGAAATATGTTTCTCAAAAAA 2824

RESULT 2
US-09-291-922-7
Sequence 7, Application US/09291922
Patent No. 6383776
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Hiltz, Bill
APPLICANT: Kinney, Tony
APPLICANT: Linney, Scott
TITLE OF INVENTION: Plant Sugar Transport Proteins
FILE REFERENCE: BB-1163
CURRENT APPLICATION NUMBER: US/09/291,922
CURRENT FILING DATE: 1999-04-14
EARLIER APPLICATION NUMBER: 60/083,044
EARLIER FILING DATE: April 24, 1998
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Microsoft Office 97
SEQ ID NO 7
LENGTH: 2601
TYPE: DNA
ORGANISM: Glycine max
US-09-291-922-7

Query Match 33.9%; Score 940.2; DB 4; Length 2601;
Best Local Similarity 66.0%; Pred. No. 1e-24;
Matches 1476; Conservative 0; Mismatches 713; Indels 48; Gaps 6;

QY 177 ATGTCGGGGGCTTTTGTGCGCATATGCGCTCCATGCGCAATCTATTTGCGGGGTGG 236
DB 175 ATGAAAGGCGCTCTGTTGTATATGCGCTCCATGCGCAATCTATTTGCGGGGTGG 234
QY 237 GACATGCAACATCGACGCTGCTTGTGTATATTAAGAAATTTCAATTCGAAAT 296

DB 225 GATATGCTACCATCGCCGGGCTATGTTATCAATTAAGAAAGCTTGTGGAA-- 292
QY 297 GAGCCACTGAGAGAGACTAATTTGTCAATGATCTTATGCGGCCACCATCTGTA 356
DB 293 ---CACTATGAAAGGCTTGTGTGGCAATGCTTGTATGAGCAACGGTATCAAC 348
QY 357 ACATCTCGGGGCTATATGAGACTGATGAGCGAGCCCTTATGTTATCTCTTCA 416
DB 349 ACATCTCGGGGCTATATGAGACTGATGAGCGAGCCCTTATGTTATCTCTTCA 408
QY 417 ATGCTGATCTTTCAGCGGCTCATGCTGATGATGCTGATGCTGATGCTGATG 476
DB 409 GTCCTATATTTCTGAGGATGTTGATGATGCTGATGCTGATGCTGATGCTGATG 468
QY 477 TTGCGACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 536
DB 469 TTGCGAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 528
QY 537 ATTTGAGAAATGAGCCCTTGTGAGATGAGGTTGCTGATATGATGATGATGATG 596
DB 529 ATATCTGAAAGGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 588
QY 597 GATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 656
DB 589 GATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 648
QY 657 CCCGATGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 716
DB 649 CCGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 708
QY 717 ACATATATTTATCTTCTGATATGATGATGATGATGATGATGATGATGATGATG 776
DB 709 ACATATATTTATCTTCTGATATGATGATGATGATGATGATGATGATGATGATG 768
QY 777 GCAAAAAGGTTGCAAAAGTTACGGGGGAAAGCAATGCTCAAGTATGTCCTT 836
DB 769 GCTAAGAGGTTGCTCAAAAGTTGCGGGAAGGAGAGATGCTCAAGTATGTCCTT 828
QY 837 CTTCTGAGAGGTTGAGAGTTGAGAGACATTTCAATTAAGAGTATCAATTTGACCT 896
DB 829 CTTCTGAGAGGTTGAGAGTTGAGAGACATTTCAATTAAGAGTATCAATTTGACCT 888
QY 897 GCCACGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 956
DB 889 GCTGACGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 948
QY 957 GGGCTGAAAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1016
DB 949 GATCCCAAGAGGCTTCTTGTATCAAAAGCTGATGATGATGATGATGATGATGATG 1007
QY 1017 AGTGTCTTCTCTGATCTGATGATGATGATGATGATGATGATGATGATGATGATG 1076
DB 1008 -----CTTGTCTCAACCATGATGATGATGATGATGATGATGATGATGATG 1056
QY 1077 GATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1130
DB 1057 GATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1116
QY 1131 GGAAGTATGAGAGACATGTTTCAAACTTGAAGATGATGATGATGATGATGATGATG 1190
DB 1117 GGAAGTATGAGAGACATGTTTCAAACTTGAAGATGATGATGATGATGATGATGATG 1176
QY 1191 CATGCCAAATGACAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1250
DB 1177 CATGCCAAATGACAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1236
QY 1251 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1310
DB 1237 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1296
QY 1311 GCAGAGGTTGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1370

```

Db      1297 ACAAGAGCCTTGAAGAAAGCTTACCTCCTCTCTTCCATGAGCATTCCTTGGCAGC 1356
Qy      1371 ATGAGAGGCA---AACCTCTTAGGGAGGGTGAAGATGTTGAGCAGCACTGATATC 1427
Db      1357 ATGAGGCGCTCACAGATGTCATGACAAAGGTCAGGTGAGCAGAGTGTGTACAGGTATAT 1416
Qy      1428 GGTGGGGGAGGAGGAGTGTGCTTGGAAGATGGTCAGAGAAGGAAGAGATGGTAGAAG 1487
Db      1417 GGTGGGCTGGGCACTGGCATGGAATGAGATGATGA---AGTGAAGATGGAACAA 1473
Qy      1488 GAAGGTGTTTCAAAAGAGTCTACTTGAACAAGAGAGGATTCCTGCTCAAGAAAGGAGC 1547
Db      1474 CAAGAGAGGTTTAAAGATTTATTAACATGAGAGAGAGTTCTGCACTGCTGTGA 1533
Qy      1548 TCAATTTGTTCACTTCCCGGTGGCCAGTGTGTTTGAAGGATGATGATTTGATCAATGCT 1607
Db      1534 TCCATTTGATCAGATTCCTCGGTGAAGGCGAA-----TTTGTCCAGGCT 1575
Qy      1608 GCTGCTTTAGTACGATCAGACCTTTCTCAAGGCTTGTGACCAAGCAGATGCA 1667
Db      1576 GCTGCTTTAGTACCAACCCGCTTTTACTCCAGAGCTTATGATGACACCCAGTT 1635
Qy      1668 GATGCTGCATGGTTCACCATCTGAGTATGCTGCCAAAGTTCAAGTTGAAGATTTG 1727
Db      1636 GGGCCCTGCAATGGTTCAACCATCTGAGACAGCTTCAAGGGGCAAGTTGAAAGCTCTT 1695
Qy      1728 TTTGAACTGAGAGTGAAGCGGTGCTGTTAGTGTGTTGGAATTCATCTTCAACAG 1787
Db      1696 CTGGAACCAAGGGGTGAAGATGATGTTGTGTGAGTGAATACAAATACCTTCAGCAG 1755
Qy      1788 TTTGCTGGAATAAAGGCTGTTCTGATCTATACCCCAAAATCTTGAGAGAGCTGTG 1847
Db      1756 TTTTCAAGGATTAATGGGTTCTATATTAACAACCTCAAAATCTTGAAGGCGGCTGT 1815
Qy      1848 GCAATTAATCTTCCAAATTTGGTCTCAGCTCGGCATCAGCATCTTGTATCACTTCT 1907
Db      1816 GAAATTTCTTTCAATATGAGCATTTGGCTCAGATTCGGCATCATTTCTTATCATGTGT 1875
Qy      1908 CTGACTACCTTAATATGCTTCTTGCAATGCTTGGCAATGCTGCTTATGATCTTTCC 1967
Db      1876 TTACACAACCTTCTGATGCTTCCCTGATAGGCGTATGAGCATGATGATGTTTCA 1935
Qy      1968 GGAAGAGGTTTTGCTGTAGGACAAATTCATTTGATGACATCTAGTATCCG 2027
Db      1936 GCGAGAGGCAAGTGTCTACTTACTACATCCCGTCTATGTTGTCTCATCTATTTTG 1995
Qy      2028 GTTGTGTCATTAATTAATTTGGGTACACTAGCCATGCTTGTCTCCACCGTCACT 2087
Db      1996 GTTCATTTGAAGCTGTGTAATTTGGCAATGTGCGCATGAGCATCTCAACAGTATG 2055
Qy      2088 GTTATCTCTACTTCTGCTGCTTCTGTTATGGAATTTGTCCATCCCAACATTTATGT 2147
Db      2056 GTTGTGTTTATTTCTGCTGCTTGTGATGGGTTATGGAACAATCCCTTTTC 2115
Qy      2148 GGAAGATCTTTCCAAACCAAGGTTGGGCTCTGATGATGCAATTTGTCCTTATATC 2207
Db      2116 TGAAGATTTTCCCACTAGGGGTCGTGCTGTGATGCTATCTGTGATGATGTTTC 2175
Qy      2208 TGGATCGAGATATCATCTCTACCTTACAGCTTCTGTGATGCTGATGATGATGAGCTG 2267
Db      2176 TGGATTTGAGACATCATCTCATCTGCTGCTGCTGATGCTGCTTATGAGACTT 2235
Qy      2268 GCGGGTGTTCAGCATATATGATGATGCTGATGCTTATTTCTTGTGTTGCTTCTT 2327
Db      2236 GGTGTGTTATTCGCAATTAACGAGTGTGTTTCTATCTCGGAGATATTTGTGTTTG 2295
Qy      2328 AAGGTCTCGAGACAAAGGGAGTCCCTTGAAGTATTAACGAATCTTTGAGATGAT 2387
Db      2296 AAGGTCTCGAGACAAAGGGAGTCCCTTGAAGTATCTGAAATTTCTTGTGGA 2355
Qy      2388 GCGAAGCAAGCGCTGC 2404
Db      2356 GCGAAGCAAGCGCTGC 2372

```

```

RESULT 3
US-09-291-922-13
; Sequence 13, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1487
; TYPE: DNA
; ORGANISM: Trifolium aestivum
US-09-291-922-13

Query Match      24.9%; Score 691.2; DB 4; Length 1487;
Best Local Similarity 80.1%; Pred. No. 3.6e-175;
Matches 852; Conservative 0; Mismatches 203; Indels 9; Gaps 3;

Qy      1395 GAGGGTGAAGATGCTGTGAGCAGCATGATATCGGTGGGAGATGACGCTTGGAAA 1454
Db      12   GAGGGTGAAGAGCAGTACAGCAGCTGTATGTGGGGGTGCAATCGCATGAAA 71
Qy      1455 TGGTCAAGAAAGAGTGAAGATGTGAAGAAAGAGGTGTTCAAAAGTCTACTTG 1514
Db      72   TGGTCAAGAGCAAGAGCGAGATGCAAGAAAGAGAGGCTTCAAAAGATCTACTTG 131
Qy      1515 CACCAAGAGGAGTCTCGCTCAAGAGGGGCTCAATTTGTTCACTTCCCGGTGTGGC 1574
Db      132   CACCAAGAGGGGTGGCGACTCAAGAGGGGCTGTGTTTCACTTCTGGTGGGGT 191
Qy      1575 GATG---TTTTGAAGGTAGTGAAGTTGTAACATGCTGCTTGTAGTACATGACCA 1631
Db      132   GATGCAAGCAAGGGGCGAGTGGTTTATACATGCTGCTTGTGTAAGCAGCTCGCT 251
Qy      1632 CTTTCTCAAAAGGCTTGTGTAACCAAGCATGTCAATGCT--GCCATGGTTACCCA 1688
Db      252   CTTTCTCAAGATCTTATGGAAGAGGTATGGGCGCCGTCAGCATATCATCA 311
Qy      1689 TCTGAGTACGCTCCAAAGTTCAAGTGTGTAAGTTGTTGAACCTGAGTGAAGGCT 1748
Db      312   TTGAGGAGAGCTCCCAAGGTTCAATCTGAAAGATCTGTTGAACCTGAGTGAAGGCT 371
Qy      1749 GCCCTGTAGTCCGAGTGTGAATTCAGATCTTCAACAGTTTGTGTAATAACGAGTGT 1808
Db      372   GATTTGTCGTCGAGTGTGAATTCAGATCTTCAACAGTTTGTGTAATAACGAGTGT 431
Qy      1809 CTGACTATACCCCAAAATTTTGAAGCAAGCTGTGTGGAGTATTTCTTCAAAATTT 1868
Db      432   CTGACTATACCTCCAAATTTGAGAGCAAGCTGTGTGGAGTATTTCTTCAAAATTT 491
Qy      1869 GGTCTACGCTGGCATGACATCCATCTTGAATCAATCTTCACTACCTTATATGCTT 1928
Db      492   GGTCTACGCTGGCATGACATCCATCTTGAATCAATCTTCACTACCTTATATGCTT 551
Qy      1929 CTTGATTTGGCTTGGCATGCTGTGATGATCTTTCGGAAGAGGTTTGTGCTA 1988
Db      552   CCAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 611
Qy      1989 GGCACATTTCAATCTTGAATGATCTTGAATCTTGTGTTGTGTTCAATTAATGAT 2048
Db      612   GGCACATTTCCATCTTGAATGATCTTGAATCTTGTGTTGTGTTCAATTAATGAT 671

```

Oy	2049	TTGGGATACATAGACCACATGCTTGGCTCCACACCGTCACTGTATGATCTACCTGCTGC	2108
Db	672	TTGAGTACGGGTGCCCCACAGCTGTGTCTTCCACAGTTAGCGTCAATGTCTTACTTCTGTCC	731
Oy	2109	TTGCTTATGAGATTGTGTCCATCCCAACATTTATGTGACAGATCTTTCACACGAG	2168
Db	732	TTTGTGATAGGGCTTTGGCCCGATCCCAACATTCATGTGACAGATTTTCCCAACGAG	791
Oy	2169	GTTGTGGCCCTCTGTATTTGCCATTTGTGCTTTACATCTTGGATCGAGATATCATCTGC	2228
Db	792	GTCGTGTGTGTCTGCATGCATATTTGGCCCTCCATCTTGTGATTTTGACATTAATGTT	851
Oy	2229	ACCTACAGCCCTCCGATGATGTGAATGCTAATTTGACGTGGGGGGTTTGACATATAT	2288
Db	852	ACCTACAGCCCTCCGATGATGTGAATGCTAATTTGATGACGGGGGTCTTTGGTATATAT	911
Oy	2289	GCAGTCGTATCTTGATTTCCCTTTGTGTCTCTTCTTAAGAGTCCCTGACACAAGGGG	2348
Db	912	GCAGTGCTTCTGTGATGCTCTTTGTGTCTCTTAACCTTAAGAGTCCCGACAGACAAAGGC	971
Oy	2349	ATGCCCTTGAAGTATTAATCCGAATCTTTTGACAGTTGTGCGAAGCAAGC--GGCTGCA	2405
Db	972	ATGCCCTTGAAGTATTAATCCGAATCTTTTGCGGTGTGGGGGAGAACGACCGCACGC	1031
Oy	2406	AAAGCTAATTTCTTTGCTACCTTTGTGTGCAACATATGCACTG 2449	
Db	1032	ATGTGCTATTCATATGAGCTTGTGTTCATAGTTTGACACTG 1075	

RESULT 4
US-09-291-922-9
; Sequence 9, Application US/09291922

```

GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Hitz, Bill
APPLICANT: Kinney, Tony
APPLICANT: Tingey, Scott
FILE OF INVENTION: Plant Sugar-Transport Proteins
FILE REFERENCE: BB-1163
CURRENT APPLICATION NUMBER: US/09/291,922
CURRENT FILING DATE: 1999-04-14
EARLIER APPLICATION NUMBER: 60/083,004
EARLIER FILING DATE: April 24, 1998
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Microsoft Office 97
SEQ ID NO 9
LENGTH: 1692
TYPE: DNA
ORGANISM: Glycine max
US-09-291,922-9

```

Query Match	21.4%;	Score 594.6;	DB 4;	Length 1692;
Best Local Similarity	66.1%;	Pred. No. 2.8e-149;		
Matches 969;	Conservative	0;	Mismatches 459;	Indels 39; Gaps 6;

Qy	929	CGGGATAGGAACAAATCAACACTTATGGGGCTGGAAGAAGCCAGTCATGGATTGGCTCG	988
Db	17	CAGAGAAAAGATCAAATTTAACTTTATAGACACAGACAGGCCACAGTCTGGGTGGCTAG	76
Qy	989	ACCTTCTAAGGAGCCCATCATGCTTTGGAAGTGTGCTTCTCTTGACATCGATCGATGGAG	1048
Db	77	ACCTGTGCTGGACCAAAATTCGTGGGC-----CTGTATCTAGGAAAGAGAG	124
Qy	1049	CATGATGAACCAAGATGTACCCCTTATGATATCCATTTGTGACACTTTTGGTAGTGTCCA	1108
Db	125	CATGGCAAAATCCAAACAGT---CTAGTGAACCTCTAGTGAACCTCTTTGGTAGTGTACA	181
Qy	1109	TGAGATATATGCTCTCAAGCTGGAGAGAATATGAGAGCACTATGTTTCCAAACTTTGGAAG	1168
Db	162	TGAGAAAGCTCCCAAGAAACAGAG-----AGCAACCTTTTTCACACTTTGGGAG	229
Qy	1169	TATGTTCAGTGTCAACAGATCAGCATGCCAAATAATAGCAGTGGATGGAAGAGATCTTCA	1228

D	b	230	TATTTTCAGTCTTGGGGAAATCAGCCAAAGAAATGAAGATTGGATATGAGAAAGCTTACG	289
O	y	1229	TAGGATGACGAGAGATACGCATCTGATGTGACAGAGGTGACTATGAGGACAATCTCCA	1288
D	b	290	CAGAGAGGTGATGATATGATCTCTGATGCT-----GGTGAATCTGATACAAATTGGCA	343
O	y	1289	TAGCCATTTGCTGTCCAGGCGAGCAAGAGTGGCGAAAGGAGAACATTGTGCACCTGG	1348
D	b	344	GAGTCCATTATCTCAGCTCAAAACMAAGATCTGGA---TAAAGACATACCTCTCTATGC	400
O	y	1349	TCACCGTAGAAGTCTTTGAGCATGAGAAGCAAAACCTTTAGGGGAGGGGTGAGATGG	1408
D	b	401	CCATATGATACCTTTGCAACGATGAGGCAAGTAGTCTTTTACATGAAATTCAGAGAAC	460
O	y	1409	TGTGAGCAGCACTGATATCGGTGGGGAGTAGGAGCTTGTGTGAAATGTTCAGAGAA	1468
D	b	461	CACGTGTAAGTACTGGAGATTGTGTGTGTGTGGCAGCTAGCATGAGAAATGCTTAAAGAG	520
O	y	1469	AGTGAGATATGTGAAAGAGAGGTGTTCAAAAGACTTACTTGCACCAAGAGGAGT	1528
D	b	521	GGGCCAGATGGAAGAGAGAGGTGCTTCAAGAGAAATATTTACACCAAGATGATGG	580
O	y	1529	TTCCGTCTCAAGAGGGGGCTCAATGTGTTACATCCCGGTGGGGAGATGTTTTTGAAGG	1588
D	b	581	TTCTGGAATCTAGAGCTGGTGTGTGGTTTCACTCC--TGGGGTGAATTTACCACTGA	637
O	y	1589	TAGTGAATTTGTATCATGCTGCTGCTTGTAGTAAAGTCAGTCAGACATTTTCTCAAGGGTCT	1648
D	b	638	CAGTGAAGTTGTACAGGCTGCTGCTGTGTGAGTCAAGCCTGCCCTTTATATAGACACT	697
O	y	1649	TGCTGAACCAAGCATGTCAAGTGTCCAGTGTTCACCCACTGAGGTAGCTGCCAAAG	1708
D	b	698	TATGCTCAACGGCCAGTGTGGAACGAGTATGATTCATCCCTGAAACAAATTCGAAAGG	757
O	y	1709	TTCAAGTTGAAAGATTTGTTTAAACCTGAGAGAGGCCGTCCCTGTACTGCTGTGG	1768
D	b	758	GCCAAAGTTGAGATCTTTTGAACCTGGGGTGAAACATGATGATGTGGGGGTGG	817
O	y	1769	AATTCAATGCTCTTCAACAGTTTGTGTGAATAAAGCGTGTCTGTACATATACCCACAAT	1828
D	b	818	AATGCAATTTCTTACGACAGTTCTCTGTGTATTAATGGGCTCTCTACATATACGCTCAAT	877
O	y	1829	TCATTAGCAAGCGTGGTGGCAAGTTATCTTCCAAATTTGTGTCACTCGGATCAGC	1888
D	b	878	TCTTGACAGGAGGTGTGGTATCTTCTTCAAGCCTTAGCCCTGTCTACTTCTTC	937
O	y	1889	ATCCATCTGTATCAGTTCTCTCACTACCTTACTATATGCTTCTTGCATTGGCTTGCAT	1948
D	b	938	ATCTTTTCTTATATGTGCGGTGACAACTGTTGATGCTTCTCTGTATPACATATGCAT	997
O	y	1949	GCTGCTTATGAGATCTTCCGGAAGAGATTTTGGCTGTGAGGACAAATTCGAATCTGAT	2008
D	b	998	GAGGCTATGATATTTTCAAGCAGAAAGAACTTTGTCTCTCAATGACATATCCCGTCTAAT	1057
O	y	2009	AGCATCTCTAAGTTATCCGTGTGTGTGTCAAATCTATATGATTTTGGGTACATGAGCCATGC	2068
D	b	1058	AGCAGCTCTTCTCATATTAAGTCCGGGAAGTCTGTGGATTTGGGATCCACTGCAAAATGC	1117
O	y	2069	TTTGTCTTCCACGCTCAGTGTATCGTCTACTTCTGCTGCTTGGTATGGAATTTGGTCC	2128
D	b	1118	ATCATCTCAACCAATTAAGTGTATATGTCTATTTCTGTTTCTTTGTCAATGGGATTTGACC	1177
O	y	2129	CATCCCAACATTTTATGTGAGAGATCTTTCACAACAGGGTGTGAGGCTCTGTATGTC	2188
D	b	1178	AATTCCTAATATATCTTTGTGCAAGATCTTCCCACTCGAGTTGTGTGTCTGTGCAATTC	1237
O	y	2189	CATTGTGCTTTACATTTCTGATCGGAGATATCATGTGCACCTCAAGCCTCTCTGTAT	2248
D	b	1238	TATTTGTGCCCTTACCTTTTGGATCTGTGATATCATTTGTCAACCTGACACAGCTCCAGTTAT	1297
O	y	2249	GCTGAATGCTATTTGACTGGCGGATGTTTTCAGCATATATGCAATGTAATGCTTGAATTC	2308

Db 1298 GGTCAATTCGTAGAGCGCTGCTGCTGTTTGGTATTATGCTGCTGCTTCAATAC 1357
Qy 2309 CTTTGTGTGCTGCTTCTTAAAGTCCCTGAGCAAGGGAGTCCCTTGAGGTATTAC 2368
Db 1358 ATGGGCTGTGCTTCTTTGAAAGTTCAGAAACCAAGGCGATGCGACTGGAGTATCAT 1417
Qy 2369 CGAATTCCTTGCACTGGTGCGCAAGCA 2395
Db 1418 TGAGTTCCTCTGTCGAGCAAAACA 1444

RESULT 5

US-09-291-922-15
; Sequence 15, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; EARLIER FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1009
; TYPE: DNA
; ORGANISM: Trifolium aestivum
US-09-291-922-15

Query Match 14.2%; Score 394.4; DB 4; Length 1009;
Best Local Similarity 74.4%; Pred. No. 8.6e-96;
Matches 497; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

Qy 1730 TGAACCTGAGTGAGGCGTGCCTGTATAGTGGTGTGAATTCAGATCCTTCAACAGTT 1789
Db 1 TGAACCTGAGTGAGGCGTGCCTGTATAGTGGTGTGAATTCAGATCCTTCAACAGTT 60
Qy 1790 TCGTGAATTAACGGGTCTCTGACTATATCCCAAAATCTTGACAGCGGTGTGCG 1849
Db 61 TCGGCTATCAAGAGTCTCTTACTACACACTCGATATCTGAACAAAGAGTGTGCG 120
Qy 1850 AGTATATCTTCCAAATTTGCTCAGCTCGGATGAGCATCCATCTTATAGTCTCT 1909
Db 121 GGTCTCTTATCAAAATTTGAGTGAATGCTCTTCTGAGCATCTATCTTATAGTCTCT 180
Qy 1910 CACTACCTTACTAATGCTTCTTGATGAGCTTGTGCGATGCTGTATGAGATCTTTCGG 1969
Db 181 GACAACTTGCTGATGCTCTCCGAGCATGCGCATGAGCATGAGATGATGTCAG 240
Qy 1970 AAGAAAGTTTGTGCTGCTGAGGACATTCCTGATGATGAGATCTCTGATGATCTGCT 2029
Db 241 AAGAAAGTTTGTGCTTCTTCAACAAATCTGCTGTGATGATGAGGCTGTGCTGTGAT 300
Qy 2030 TGTGTCATCTAATTAATGATTTGGGTACACTAGCCCATGCTTGTCTCCACGCTAGTGT 2089
Db 301 TTTAGGAAATGCTTGTGATGTGGAACATGTCAGCGCTGCTCTCAACATCAAGCT 360
Qy 2090 TATGCTACTTCTGCTGCTGCTGATTAAGGATTTGATGCCATCCCAAAATTTATGTGC 2149
Db 361 CATGCTATTTCTGCTGCTGCTGATGAGGCTTGGGCTATCCCAAAATTTCTGCGCC 420
Qy 2150 AAGATCTTTCACCAAGGCTTGTGCGCTCTGATGATGCTATTTGCTTATCATTTG 2209
Db 421 GAGATTTTTCGCCACTCTGTGCTGAGCATCTGATGAGCATCTGCGCGCTTACCTTCTG 480
Qy 2210 GATCGAGATATCATGTCACCTAGACGCTTCTGATGATGATGATGATGATGATGATG 2269
Db 481 GATCGAGATATCATGTCACCTAGACGCTTCTGATGATGATGATGATGATGATGATG 540

Qy 2270 GGTGTTTTCAGCATATATGACATGATGCTGATTTCTTGTGCTGCTTCTTCA 2329
Db 541 TGAAGTCTTCCGCAATATATGCGATGTTGTGATGCTACCTTTGATTTGCTTACATGA 600
Qy 2330 GGTCCCTGAGCAAGGAGATGCGCTTGAAGTATTAACGAATCTTGTGAGTGTGTC 2389
Db 601 GGTCCCTGAGCAAGGAGATGCGCTTGAAGTATTAACGAATCTTGTGAGTGTGTC 660
Qy 2390 GAAGCAAG 2397
Db 661 AAGCAAG 668

RESULT 6

US-09-291-922-5
; Sequence 5, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; EARLIER FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 870
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-291-922-5

Query Match 8.2%; Score 227; DB 4; Length 870;
Best Local Similarity 74.9%; Pred. No. 5.3e-51;
Matches 284; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

Qy 2020 TTATCCTGTTGTGCTCAATCTAATTTGATTTGGGTACAGCCCATGCTTGTCTCCA 2079
Db 12 TTAACCTTATTTGCTCAATATTTGATGTGGGACATGTTATGCTTACTGCTCA 71
Qy 2080 CCGTCAAGTATATGCTTACTTCTGCTGCTGCTGATGAGATTTGCTCCATCCCAACA 2139
Db 72 CAGTCAAGTATATGCTTACTTCTGCTGCTGCTGATGAGATTTGCTCCATCCCAACA 131
Qy 2140 TTTTATGTGCAAGATCTTTTCCAAACAGGCTTGTGCTGCTGATTTGCTGCT 2199
Db 132 TTTCTGTGCAAGATTTTCCGACACCGTTCGAGATCTGATGAGCATCTGTGCCC 191
Qy 2200 TTACATTTGATGAGGATATCATGTCACCTACAGCTTCTGATGATGATGATGAT 2259
Db 192 TTAATTTGATGAGGATATCATGTCACCTACAGCTTCTGATGATGATGATGAT 251
Qy 2260 TTGAATGAGGATTTTTCAGATATATGAGTGTGATTTCTTTGTGCTG 2319
Db 252 TTGAATGAGGATTTTTCAGATATATGAGTGTGATTTCTTTGTGCTG 311
Qy 2320 TTTTCTTATGATCCCTGAGACAAAGGAGTCCCTTGAAGTATTTACCAATCTTTG 2379
Db 312 TTTTCTTATGAGTGTGAGGACAAAGGAGTCCCTTGAAGTATTTACCAATCTTTG 371
Qy 2380 GAGTGTGCGAAGCAAG 2398
Db 372 CTGTGAGCAAGCAAGC 390

RESULT 7

US-09-313-294A-4597
; Sequence 4597, Application US/09313294A

```
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalquidi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 4597
; LENGTH: 155
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Inceyte ID No. 6476212 700348695H1
; NAME/KEY: unsure
; LOCATION: 45, 56, 77, 83, 140, 154
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-4597
```

```
Query Match          5.2%; Score 143.4; DB 4; Length 155;
Best Local Similarity 96.0%; Pred. No. 5.5e-29;
Matches 144; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1039 GTCATGGGAGCATGTCGAACCAAGTGTACCCCTTATGATCCGATTGGACACTTTTG 1098
```

```
DB 1 GTCATGGGAGCATGTCGAACCAAGTGTACCCCTTATGATCCGATTGGACACTTTTG 60
```

```
QY 1099 GTAGTGTCCATGAGATATGCTCAAGCTGAGAGAGATAGAGAGACATTTTCCAA 1158
```

```
DB 61 GTAGTGTCCATGAGATATGCTCAAGCTGAGAGAGATAGAGAGACATTTTCCAA 120
```

```
QY 1159 ACTTTGGAAGTATGTTCACTGTCACAGATC 1188
```

```
DB 121 ACTTTGGAAGTATGTTCACTGTCACAGATC 150
```

```
RESULT 8
US-09-291-922-11
; Sequence 11, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (421)
; NAME/KEY: unsure
; LOCATION: (434)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (441)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (458)
```

```
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (483)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (493)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (498)
US-09-291-922-11
```

```
Query Match          4.4%; Score 123.2; DB 4; Length 510;
Best Local Similarity 72.7%; Pred. No. 2.5e-23;
Matches 184; Conservative 0; Mismatches 66; Indels 3; Gaps 2;
```

```
QY 168 GAGGCAAGATGTCGGGGGCTGTTTGTCCGCAATAGCGCTCCATGCGCAATTTATG 227
```

```
DB 194 GGGAGGAAGATGTCGGGGGCTGCTGCACTGTGCGATGGGGCTTCATTTGCAATTTGCTG 253
```

```
QY 228 CAGGGGTGGGCAATGACCAATGCGACCTGCTGTTCTGTATATTAAGAGAAATTTCAA 287
```

```
DB 254 CAGGGGTGGGCAATGACCAATGCGACCTGCTGTTCTGTATATTAAGAGAAATTTCAA 313
```

```
QY 288 TTGCAAAATGAGCCCACTGTGAGGAGACTAATGTGTCATGTCACTTATGCGGCCAAC 347
```

```
DB 314 CTCGAAATTAATCCGACTGTGAGGAGGCTCATCGTGGCA--TGCTCATGCGGGTGAAC 371
```

```
QY 348 ATCGTTACTACTTCTCCGGGCACTTATACAGACTGATTTGGCCGACGCTTATTT 407
```

```
DB 372 ATCATCAC-ACATTTCTCGGGGCAAGTATCAATGGGTTGCGGGCCCTTANCCATCTCC 430
```

```
QY 408 CTCCTCTCAATTC 420
```

```
DB 431 TTGATTTCAATTC 443
```

```
RESULT 9
```

```
US-09-291-922-3
; Sequence 3, Application US/09291922
```

```
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 443
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (193)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (388)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (435)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (439)
US-09-291-922-3
```

```
Query Match          4.1%; Score 114.2; DB 4; Length 443;
Best Local Similarity 74.5%; Pred. No. 6.1e-21;
```

Matches 143; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 177 ATGTGGGGGCTGTTCTTGTGCGCATATGCGCTCATGCGCAATCTATGAGGGGTG 236
DB 205 ATGGCGGGGCGGTGCTGCTGCGCATGCGGCTCCATCGGCAATCTTGCGAGGGCTG 264

QY 237 GACAAAGCCACCATGCGAGTGTCTTCTATATTAAGAAAGAAATTTCAATTGCAAAAT 296
DB 265 GATTAATGCAACCATTTGAGGTGGTACTGATCATCAAGAAAGAAATTTCAATTGCAAAAT 324

QY 297 GAGCCCACTGTGAGGAGCACTAATTTGTCAATGTCACTTATCGGCCCAACATCTTACT 356
DB 325 GACCCCTTATGAAAGTGTGATGCTGCGCATCTCCCTCATTTGGGGCAACATCAACAG 384

QY 357 ACATCTCCGG 368
DB 385 ACGTCTCTGCG 396

RESULT 10
US-09-291-922-23
; Sequence 23, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Linney, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; EARLIER FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 23
; LENGTH: 1853
; TYPE: DNA
; ORGANISM: Glycine max
US-09-291-922-23

Query Match 3.5%; Score 96.2; DB 4; Length 1853;
Best Local Similarity 48.6%; Pred. No. 8.1e-16;
Matches 297; Conservative 0; Mismatches 308; Indels 6; Gaps 1;

QY 186 GGTGTTCTGTGCGCATATGCGCTCATGCGCAATCTATGAGGGGTGAGCAATGCC 245
DB 257 GCTTTGCTGTGCTATGCTGCGCTCATGCTTCCATCTTGTGTTATGATATGGA 316

QY 246 ACCATGCAAGTGTCTGTTCTGTAATTAAGAAAGAAATTTCAATTGCAAAATGAGCCACT 305
DB 317 GTGATAGTGGAGCAACCATATATCAATAAAGGAACTGAAGTCTGAGACGCAAAATC 376

QY 306 GTGGAGGAGCACTATTTGTCAATGTCATTTATGAGGCGCAACATCTTACTACATTTCC 365
DB 377 GAG-----ATCTGTGCGAATCAACCTATTAATCTCTGATAGGCTCATGCTTCTGCGC 430

QY 366 GGGCCATTAATGAGCTGATGAGGCGAGCGCCATGCTTATTTCTCTTCAATTTGTGAC 425
DB 431 GCGAGAACTTCCAGCTGATAGTCCCGTTACACATTTGTTTCCCGGCAACATCTTC 490

QY 426 TTCTTCAAGGCGCTCATGCTATGATGCTCTTAATGCTATGCTGCTGTTGGACGC 485
DB 491 TTGTGCGAGCACTTCTCATGAGGTTTCTCCCAATTAATCTTCTCATGTTTGGCGGT 550

QY 486 TTCTGATGATGATTTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 545
DB 551 TTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 610

QY 546 ATAGCCCTTTCGAGATTAAGGTTGCTGATACACTACCAATTCAGTGGATCAGA 605
DB 611 GTCTCCCGGCTCTCTGCTGCTGCT 670

QY 606 GGAATGTTCTGTCTACTGATGCTGTTGGAGATGCTCTGTGCGCATACACCGATTGG 665
DB 671 GGGATTAATTAATTTGATATATATCAATCAATGATTTTGAAGCTGACATAAGGTGGA 730

QY 666 AGAATTAATGCTTGGTGTGCTGCGCATACCTTATGTTCTTCTTGTGTTGCAATATTT 725
DB 731 TGGCAATGATGCTTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 790

QY 726 TATCTTCTGATCTCAAGATGCTGCTTGAAGAGGTGATGAGGAGGCAAGAAAAAG 785
DB 791 GCGATGCGGAGTCCCAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 850

QY 786 GTGTGCAAAA 796
DB 851 GTGCTTAACAA 861

RESULT 11
US-09-291-922-21
; Sequence 21, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Linney, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; EARLIER FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-291-922-21

Query Match 3.0%; Score 84; DB 4; Length 2017;
Best Local Similarity 47.1%; Pred. No. 1.6e-12;
Matches 323; Conservative 1; Mismatches 356; Indels 6; Gaps 2;

QY 112 CTCTTCCGCTGCTGCTGCTGCTTCTTCAAGAGCGGTGACCTCGGACGATATCTTGAAG 171
DB 82 CTGACCAACCGAATGCTTCCGCGCTGCGGAGCGGTGCGCGCAAGAAAG 141

QY 172 ACAATATGCGGGCTGCTTGTGCGCATAGTCCCTCATGCGCAATCTATGCAAG 231
DB 142 GCMAAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 201

QY 232 GGTGGACCAATGCCACATGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 291
DB 202 GCTACGATATGCGGTATGAGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 259

QY 292 AAAATGAGCCCACTGTGAGGAGCACTAATTTGTCAATGCTCAATTCGCGCCACCATCG 351
DB 260 -CAGTGAAGGAGGTGAGGTGCTCATGCGGATGCTGAGCC--TCACTGCTCATCG 315

QY 352 TTACTACATTTCTCGGGCAATTTATGAGCTGATTTGCGGAGCGCCATATGCTTATCTCT 411
DB 316 GCTCTTCCGCGGCGGAGGAGCTGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 375

QY 412 CTTCATTTCTGATCTTCTTCAAGCGGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 471
DB 376 CCGCGCTATATTTCTTCCGCGGAGGCTTCTCATGAGGCTTGCCTCACTACGCAATGC 435

QY 472 TGTCTTGGCAAGCTTCTGATGATGATTTGATTTGCTTGGCTGCTGCTGCTGCTGCTGCTGCT 531
DB 436 TCAATTTGCGGCGCTTCTGCGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 495

QY 532 TGTACATTTAGAAATAGCCCTTCGAGATTAGAGTTGTGTAATACATACCAACAT 591
| | | | |
Db 496 TGTACACCGCCGAGGTGTGCGCGGTGGCGGTGCTTCTGACGTGTCCCGAG 555
QY 592 TCAGTGCATCAGAGAAATGTTCTTGTCTACTGATGTTGGAGATGTCCTGTGCG 651
| | | | |
Db 556 TGTTCATCAACTTCGCGATCCTGCTCGAGTACGTCTCGAATAGTTCTCTCCGTTGC 615
QY 652 CATCACCGGATTTGAGAAATATGTTGTTGCTGCGCATACCTTCATTTCTTCTTG 711
| | | | |
Db 616 CGCTGACCTCGGGGTGGCGCATATGCTGCGATCGCGCGCGCGCTGCTGCTGCTG 675
QY 712 GTTTGACATATTTATCTTCTGTAATCTCAAGATGCTGTTGCAAAAGTGGATG 771
| | | | |
Db 676 CGCTATGTTGCTCGCATGCGGAGTGGCGGCGGTGCTGATGATGAAGGACGCTG 735
QY 772 CAGAGGCAAAAAGTTGTTGCAAAAG 797
| | | | |
Db 736 CGAGCGCCAAAGGTGTGCTGAGAGAG 761

RESULT 12

US-09-291-922-25
; Sequence 25, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hiltz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Kinney, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291, 922
; EARLIER APPLICATION NUMBER: 60/083, 044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 25
; LENGTH: 2089
; TYPE: DNA
; ORGANISM: Trifolium aestivum
US-09-291-922-25

Query Match 2.8%; Score 78.2; DB 4; Length 2089;

Best Local Similarity 48.2%; Pred. No. 5.7e-11;
Matches 317; Conservative 0; Mismatches 328; Indels 12; Gaps 3;

QY 144 CCGGTGACCTCGAGCATATCTTGAGACAAAGATGTCGGGGCTGTTCTTGCGCCATA 203
| | | | |
Db 95 CCGGGCGCGCTGAGGCCAAGAAAGGCGCAACGTGAGTTGCTTCCCTGCGCCATC 154
QY 204 GTGCGCTTCATCGGCAATCTATTTGAGGGGTGGACATGCCACATCGCAGTGTCTT 263
| | | | |
Db 155 CTCGCTTCATGACCTCCATCTCTCGGCTACGATCGGGGTATGAGCGGAGGTG 214
QY 264 CTGTATATAAAGAAATTTCAATTGCAAAATGAGCCACTGTGAGGAGCTAATTGG 323
| | | | |
Db 215 CTGTACATCAGAAAG-----ATCGAAGATCMAAGAACCCAGCTGAGGCTCCATG 268
QY 324 TCAATGTCACTTATCGGGCGCACCATGTTACTACATTTCTCGGGCGCATTTACAGCTG 383
| | | | |
Db 269 GGCATCTCAAGCTGATCTGCTCATTTGCTTTCATATCTGACTTTTCAAGCGCTCATC 443
QY 384 ATTGCGCAGCCCTATGTTATCTCTTTCATATCTGACTTTTCAAGCGCTCATC 443
| | | | |
Db 329 ATCGGCGCGGCTTACCAATGCTTTCGCGCGCTCATCTTCTTTCGCGGCGCTCATC 388
QY 444 ATGCTATGCTCTCTAATGCTATGCTGCTGTTGGACGCTCTGTAATGATTTGGT 503
| | | | |
Db 389 ATGGGCTTCTCGCTCACTAAGCCAGCTCATGTTGGGCGCTTCTGAGCGGCATCGGC 448
QY 504 ATTGCTTGGCTGTCAAGCTTGTGCTTTGTACATTTAGAAATAGCCCTTCGAGATT 563

Db 449 GTGGGTAGCTCTCATATGCGCGCCGCTGAACAGGGCGAGGTGTCCCGCGCTGCC 508
| | | | |
QY 564 AGAGTTTCTGTAATACACTACCAATTCAGTGAATAGAGAAATGTTCTGTCAATC 623
| | | | |
Db 509 CGTGGGTTCTCAATCTTCCCGGAGGTGTTCAATCACTTCGGATCTCTCTCGAAT 568
QY 624 TGCATGTGTTTGGAGTGC---CCTGTGCGCATCACCCGATGAGAAATATGCTGAT 680
| | | | |
Db 569 GTCTCAACTTCGCTTGGCGCGCTTCTCTCGGCTGCGCTGCGCATTTATGCTCGGC 628
QY 681 GTGCTGGGATACCTTCATTTGTTCTTCTTGTGTTGCAATATTTATCTTCTGAACT 740
| | | | |
Db 629 ATAGCGCGGTCGCTC---CGTCTGCTCGGCTTCATGTTGCTGCGCATGCGGATGCT 685
QY 741 CCAAGATGCTCTGTTAGCAAAAGTGTGATGTCAGAGGCAAAAAGTTGTTGCAAAAG 797
| | | | |
Db 686 CCGCGTGGCTCTCATGAAAGGCGCTCTCGCGAGCCAAAGTTGTTGCTTGCAGAG 742

RESULT 13

US-08-232-463-14/C
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHREIFINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEEX: 899149

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgp-F18
US-08-232-463-14

Query Match 2.5%; Score 69; DB 1; Length 7218;

Best Local Similarity 5.3%; Pred. No. 3e-08;
Matches 21; Conservative 227; Mismatches 147; Indels 0; Gaps 0;

```
QY 1152 TTTCAGAACTTGGAGTATGTCAGTGTACAGATCCAGCAAAAATGAGCAGTGG 1211
DB 1479 TTACCTATCTAGCAAGTATGTAAGATAGAGATTGTCACRRRRRRRRRRRR 1420
QY 1212 GATGAGAGAAATCTTCATAGGATGACGAGTACGATCTGATGTCAGAGAGTAC 1271
DB 1419 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1360
QY 1272 TATGAGAGCAATCTCATAGCCATTCTGTCCAGCAGCAACAGTCCGAGGAG 1331
DB 1359 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1300
QY 1332 GAGATTGTCCAGCATGTCACGTCGTGATGATGATGAGGAAACCCCTTA 1391
DB 1299 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1240
QY 1392 GCGGAGGATGATGATGTCAGCAGCTGATATGTCGGGAGTGCAGCTTGCTTG 1451
DB 1239 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1180
QY 1452 AAATGTCAGAGAGAGAGTGAATGTAGAAAGAGTGTCTTCAAAAGAGTCTAC 1511
DB 1179 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1120
QY 1512 TTGCAACAGAGAGGAGTCTCTGCTCAAGAGGG 1546
DB 1119 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1085
```

RESULT 14
US-09-291-922-27
Sequence 27, Application US/09291922

Patent No. 6383776
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Hiltz, Bill
APPLICANT: Kinney, Tony
APPLICANT: Tingey, Scott
TITLE OF INVENTION: Plant Sugar Transport Proteins
FILE REFERENCE: BB-1163
CURRENT APPLICATION NUMBER: US/09/291,922
CURRENT FILING DATE: 1999-04-14
EARLIER APPLICATION NUMBER: 60/083,044
EARLIER FILING DATE: April 24, 1998
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Microsoft Office 97
SEQ ID NO 27
LENGTH: 1872
TYPE: DNA
ORGANISM: Trillium aestivum
US-09-291-922-27

Query Match 2.4%; Score 65.8; DB 4; Length 1872;
Best Local Similarity 44.7%; Pred. No. 1.1e-07;
Matches 302; Conservative 0; Mismatches 367; Indels 6; Gaps 1;

```
QY 123 GCTTGTCTTCTTACAGAGCCGCTGACCTGCGAGCATATCTTGAGAGCAAGATGTC 182
DB 122 GCTTGTCTGCTGCTCCGAGCGGAGGAGTCAATCCAAAGAAAGGCAATTTCAAG 181
QY 183 GGGGCTGTCTGTGCGCCATAGCGCTCATCGGCAATCTATGAGGGGTGGCAGT 242
DB 182 TAGGCTTACCTGCGCCCTGTGTGCTTCATAGGCCACATGCTCTGAGTACGAGT 241
QY 243 GCCACCATGCACTGCTGTCTGTATATAAGAAATTTCAATTTGCAAAATGAGCC 302
DB 242 GGGGATAGGCGGTGCTGCTGTATCAATCAAGAGGACCTGAGATCAAGAGCGT 299
QY 303 ACTGAGAGGAGCAATATGTGTCAATGTCACTTATGCGGCCCAATGCTTATCAATTC 362
DB 300 ---AGCTGAGATCATGTAGTGGGATCTGAGCGGTGATGCGGCTCATGGGCTCTTC 355
QY 363 TCGGGGCAATATCAAGTCTGATGGCGGAGCCCTATGCTTATCTCTTCAATTCG 422
```

```
DB 356 GCGCGAGAGAGCTCCGATCTGGGTGCGCCGCGGCTGACCGTCTGTCGCGCCGCA 415
QY 423 TACTTCTTACGCGGCTCATATGATGATGATGATGATGATGATGATGATGATGAT 482
DB 416 TTCAACAAGGCTCTTCTCTCATGGCTTCCGCGTCACTAGCCATCTCATGGTGG 475
QY 483 CGCTTCATGATGATTTGATTTGGCTTGGCTGACAGCTTGGCTTGGCTTGGCTTGG 542
DB 476 CGCTTCATGATGATTTGATTTGGCTTGGCTGACAGCTTGGCTTGGCTTGGCTTGG 535
QY 543 GAATATGAGCCCTTGGAGATTAAGGTTGCTGAATACATCAACCAATTCAGTATCA 602
DB 536 GAGGTGTCCCGGCTGCGCCGCGGCTTCTCTCACTTTCACGAGGTGATCAT 595
QY 603 GAGGAGATGTTCTTGTATCTGATGATGATGATGATGATGATGATGATGATGAT 662
DB 596 GTGGGATCTCTCTTGGCTAGCTGCTCAACTAGGCTTGGCGGCTCCGCTCACCTC 655
QY 663 TGAAGATTTATGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 722
DB 656 AGCTGCGCGTATGCTGCGATGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 715
QY 723 TTTTATCTTCTGATCTCAAGATGCTGCTGTAAGCAAGTGGATGGAGGCAAAA 782
DB 716 TTGCGATGCGGAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 775
QY 783 AAGTGTTCAGAAAG 797
DB 776 GCGGTCTGCGCAAG 790
```

RESULT 15
US-09-291-922-19
Sequence 19, Application US/09291922

Patent No. 6383776
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Hiltz, Bill
APPLICANT: Kinney, Tony
APPLICANT: Tingey, Scott
TITLE OF INVENTION: Plant Sugar Transport Proteins
FILE REFERENCE: BB-1163
CURRENT APPLICATION NUMBER: US/09/291,922
CURRENT FILING DATE: 1999-04-14
EARLIER APPLICATION NUMBER: 60/083,044
EARLIER FILING DATE: April 24, 1998
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Microsoft Office 97
SEQ ID NO 19
LENGTH: 1914
TYPE: DNA
ORGANISM: Zea mays
US-09-291-922-19

Query Match 2.3%; Score 63.4; DB 4; Length 1914;
Best Local Similarity 45.0%; Pred. No. 5e-07;
Matches 284; Conservative 0; Mismatches 341; Indels 6; Gaps 1;

```
QY 167 GGAGAGCAAGATGTCGGGGGCTGTTCTGTGCGCATATGCTTCATGCGCAATCTATT 226
DB 97 GAGGAGCAAGTCAAGTATGCTTCATATGTCATCTGCTGCTTCATAGGCTCTGTAT 156
QY 227 GCGAGGAGTGGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 286
DB 157 CTTGCTATATGATTTGGGATGATGAGCGGCGCATGATCAAGAGAGAGCTGAA 216
QY 287 ATTGCAAAATGAGCCCATGTCGAGAGGACTAATTTGTCAATGATCACTTATGGGCGAC 346
DB 217 TATCAAGAGCGTGC-----AGCTGAGATCTGATGCGGATCTCACTTCTTACTGCT 270
QY 347 CATGTTACTATCTCTCGGGCCATTTATCAAGTCTGATGGCGGAGCCCTATGCTTAT 406
```

```
Db 271 GTTCGATCCTTCGCTGCGCGCGGACGTCGACAGGATCGGGCGCCGCTTGACCGTCGT 330
Qy 407 TCTCTCTCAATTCCTGTAATCTTCATGAGGGCCCTCATGCTATGATGCTCTAATGTCGA 466
Db 331 GTTCGCGCGTGCATCTTCTTCGTGGGCTGCTGCTCATGGGTTGCGCGTCACTACGG 390
Qy 467 TGTCTGCTGTTGGCAAGCTTCGTAGATGATTTGGTATTTGGCTTGGCTGTACCGCTGT 526
Db 391 CATGCTCATGGGGCGCGCGCTTCGTGCGCGAGTCGAGTGGGCTACGGGGGCATGATCGC 450
Qy 527 GCTTTGTACATTTAGAAATAGCCCTTCGAGATTAAGGTTGCTGAATACACTACC 586
Db 451 GCCCGTGTACAGCGCGAGATCTCGCTGCGCGCGTCCCGTGGCTTCCTGACCACTTCCC 510
Qy 587 ACAATTCAGTGAATCAGAGAAATGTTCTGTATACATGATGATGTTGGATGTCCT 646
Db 511 GAGGCTGTTCAACAACATCGCATCTGCTTGGCTACCTGTCGAACCTTGCGCGCG 570
Qy 647 GTGCGCATCACCAGATTGAGAAATTATGCTGTGCTGCGGATACCTTCATGTTCTT 706
Db 571 CTTCCCGCTCCAGCTCGGCTGGCGGCTCATGCTCGGCATTTGGGCGAGTTCCGTCGGGCT 630
Qy 707 CTTTGGTTGACATATTTTATCTTCCTGAATCTCCAGATGAGCTGTTAGCAAGGTCG 766
Db 631 GCTCGGCTCTCTGTGTTCTGTCATGCCCCGAGTCGCTCGGTGAGTCTTGAAGGGCCG 690
Qy 767 GATGCGAGGCAAAAAAGGTGTGCAAG 797
Db 691 CTTGCGGAGCGCCAGGGCTGTGCTAGAGAG 721
```

Search completed: January 5, 2004, 15:50:46
Job time : 175 secs

THIS PAGE BLANK (USPTO)